4-1-2014

Molecular Ecology of Globally Distributed Sharks

Christine B. Testerman
Nova Southeastern University Oceanographic Center, powaser@nova.edu

NOVA SOUTHEASTERN UNIVERSITY OCEANOGRAPHIC CENTER

Molecular ecology of globally distributed sharks

By

Christine B. Testerman

Submitted to the Faculty of
Nova Southeastern University Oceanographic Center
in partial fulfillment of the requirements for
the degree of Doctorate of Philosophy

Nova Southeastern University

April 2014

Dissertation of
Christine B. Testerman

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctorate of Philosophy:

Marine Biology

Nova Southeastern University
Oceanographic Center

April 2014

Approved:

Dissertation Committee

Major Professor :_____
Mahmood Shivji, Ph.D.

Committee Member :_____
George Duncan, Ph.D.

Committee Member :_____
Dave Gilliam, Ph.D.

Committee Member :_____
Jose Lopez, Ph.D.

Table of Contents

General Abstract

Many sharks have life history characteristics (e.g., slow growth, late age at maturity, low fecundity, and long gestation periods) that make their populations vulnerable to collapse due to overfishing. The porbeagle (*Lamna nasus*), bull shark (*Carcharhinus leucas*), great hammerhead (*Sphyrna mokarran*), and smooth hammerhead (*S. zygaena*), are all commercially exploited. The population genetic structure of these species was assessed based on globally distributed sample sets using mitochondrial control region (mtCR) sequences and/or nuclear markers. Complex patterns of evolutionary and demographic history were inferred using coalescent and statistical moment-based methods. All four species showed statistically significant genetic partitioning on large scales, i.e., between hemispheres (*L. nasus* mtCR ϕ_{ST} = 0.8273) or oceanic basins (*C. leucas* nuclear F_{ST} = 0.1564; *S. mokarran* mtCR ϕ_{ST} = 0.8745, nuclear F_{ST} = 0.1113; *S. zygaena* mtCR ϕ_{ST} = 0.8159, nuclear F_{ST} = 0.0495). Furthermore, *S. zygaena* mtCR sequences indicated statistically significant matrilineal genetic structuring within oceanic basins, but no intra-basin structure was detected with nuclear microsatellites. *S. mokarran* showed statistically significant genetic structure between oceanic basins with both nuclear and mitochondrial data, albeit with some differences between the two marker types in fine scale patterns involving northern Indian Ocean samples. A microsatellite assessment of *C. leucas* demonstrated no population structuring within the Atlantic or Indo-Pacific, with the exception that samples from Fiji were differentiated from the remaining Indo-Pacific Ocean locations. In contrast, the *L. nasus* mitochondrial and nuclear ITS2 sequences revealed strong northern vs. southern hemispheric population differentiation, but no differentiation within these hemispheres. These geographic patterns of genetic structure were used to determine the source of fins obtained from the international fin trade and to develop forensic tools for conservation.

Key words: molecular ecology, population genetic structure, conservation, management, philopatry, microsatellites, mitochondrial control region, differentiation, diversity, sharks

General Introduction

Many marine animals are highly vagile during one or more life stages. Some large-bodied species (e.g., mammals, birds, turtles, sharks, and billfish) frequently travel great distances as adults, whereas many others (e.g., zooplankton and early life stages of broadcast spawning species) can disperse long distances, typically floating passively on oceanic currents. In addition, there are relatively few obvious physical barriers in the oceans other than very large distances, currents, and continental landmasses. This combination led to a general, although largely untested, assumption that most marine species were panmictic with high geographic connectivity and little if any genetic population structure, or patterns of genetic differentiation, across vast distances. Large-bodied sharks, many of which have global distributions and are capable of high vagility as adults, have been assumed to fit the genetically panmictic model of population structure worldwide. However, in contrast to this traditional view, recent genetic studies utilizing high-resolution markers are uncovering varying degrees of genetic population structure within and between oceanic basins in many shark species. This population division has been attributed to biogeographic barriers to movements and/or behavioral factors. Examples of biogeographic barriers include the Indonesian throughflow current (Dudgeon *et al.* 2009), the equatorial current (Mendonça *et al.* 2011) and the IndoPacific and eastern Pacific biogeographic barriers (Whitney *et al.* 2012). The most often suggested behavioral mechanism is female philopatry and widespread male-mediated gene flow (Pardini *et al.* 2001; Keeney *et al.* 2005; Blower *et al.* 2012; Daly-Engel *et al.* 2012; Sodré *et al.* 2012; Mourier & Planes 2013). A second behavioral factor that can also result in significant population structure is restricted dispersal tendencies as exhibited by some sharks (Gaida 1997; Karl *et al.* 2012).

Many elasmobranchs are impacted by commercial and recreational fisheries, both as targeted species and as bycatch. There have been several reports of global and regional declines in elasmobranch abundance due primarily to the effects of overfishing and habitat destruction (Baum & Myers 2004; Baum *et al.* 2005; Baum & Blanchard 2010). Sharks are considered to be top predators in marine food webs that typically have at least four trophic levels (Cortes 1999). There is increasing evidence that direct and indirect effects of fishing are affecting the composition and diversity of elasmobranch

and total fish assemblages through trophic interactions, contributing to the overall degradation of coral reef ecosystems by decreasing herbivorous fishes and thus shifting the reefs from coral to algal dominated environments (Shepherd & Myers 2005; Myers *et al.* 2007; Ferretti *et al.* 2010). Additionally, most elasmobranchs have K-selected life history traits, i.e., slow growth, late age at maturity, low fecundity, long gestation periods, and a long life span (Calliet *et al.* 2005). These characteristics result in a reproductive potential that is more similar to large mammals than teleost fishes, and a direct relationship between juvenile stock size and recruitment to the adult population (Musick *et al.* 2000). Thus elasmobranch populations have lower intrinsic rates of population increase, rendering them more susceptible to population collapse, extirpation or even extinction due to overfishing, habitat loss or degradation, and other anthropogenic effects (Walker 1998; Castro *et al.* 1999; Musick *et al.* 2000; Au *et al.* 2009). In addition, coastal and estuarine elasmobranch species are more likely to be impacted by human activities than are benthic, oceanic, or pelagic species that have a larger spatial buffer zone. These issues coupled with concerns about continued elasmobranch declines worldwide have resulted in urgent calls for improved elasmobranch conservation and management efforts based on a better understanding of their population dynamics.

Knowledge of elasmobranch species' population genetic structure, dispersal patterns, and geographic philopatric behavior can have important implications for fisheries management and conservation strategies. Genetic analyses using suitable markers from both mitochondrial and nuclear genomes can provide insight into several aspects of the population dynamics of a species, including identification of discrete populations, management units, and evolutionarily significant units; estimates of migration and gene flow among populations; assignment of individuals or body parts of unknown provenance to known species or populations; indications of whether a population is growing or declining with an inference of past demography; and estimates of lineage divergence times (Pearse & Crandall 2004; Excoffier & Heckel 2006).

Elasmobranchs are also an interesting group in a broad scientific sense. There are more than 1200 extant elasmobranch species (Naylor *et al.* 2012), yet despite their importance in fisheries, and their often meso- to apex-predator roles in a wide range of

ecosystems, elasmobranch life history parameters and population dynamics remain poorly described. The elasmobranch fossil record stretches back approximately 400 million years, although modern sharks and rays appeared about 200 million years ago. Elasmobranchs also exhibit an extensive range of life history traits and reproductive modes. They occupy the entire spectrum of marine ecosystems from estuaries and inshore coastal habitats to benthic and open ocean habitats, with individual species superbly adapted to their particular environmental niche. Many species, especially the large sharks, have global distributions, although many smaller elasmobranchs have relatively restricted or disjunct distributions. What little is known about chondrichthyan (cartilaginous fish) genomes has raised intriguing evolutionary questions. Apparently this lineage did not undergo the complete genome duplication event evident in teleost fishes (Donoghue & Purnell 2005). This genome duplication may have resulted in an accelerated rate of evolution in teleost fish and may partially explain the greater architectural similarity of mammalian and elasmobranch genomes (Venkatesh *et al.* 2007; Naylor & Aschliman 2013). This surprising similarity in genome organization and the position of elasmobranchs as basal gnathostomes (jawed vertebrates) means that elasmobranchs will play an important role in future comparative vertebrate genomic studies. A solid understanding of the molecular ecology of this diverse group will be informative for such studies.

To help further this understanding, four shark species were selected as the subjects of a molecular ecology investigation focused on population dynamics and conservation issues. All four species are of conservation concern and share the following characteristics: 1) They have cosmopolitan global distributions which facilitated evaluation of global patterns of genetic diversity, regional differentiation, and potential barriers to gene flow; 2) They are taken in recreational and commercial fisheries, both as bycatch and as targeted species; 3) They have K-selected life histories making them strongly vulnerable to anthropogenic pressures; and 4) They are capable of long distance movements as demonstrated by conventional tag/recapture data. The study species and associated habitat use and conservation status are provided in Table 1. Additional information about each of these species is presented within its respective chapter.

The main objective for this dissertation was to use mitochondrial DNA and nuclear microsatellite data to assess the global population genetic structure of each of the study species, including lineage and species delineation, genetic diversity, and evolutionary and demographic history. The combined use of mitochondrial and nuclear markers accounted for female and male mediated gene flow, respectively, and enabled inferences of population history on contemporary and historic timescales. A secondary objective was the application of this genetic information to conservation and management issues by using genetic methods to determine the geographic source of market fins obtained from the international fin trade and developing forensic tools for conservation.

Table 1. Study Species. Conservation status includes the IUCN Redlist assessment (the International Union for the Conservation of Nature) and, when applicable, listing on CITES (Convention on International Trade in Endangered Species of Wild Flora and Fauna).

Species	**Habitat**	**Conservation Status**	**Fisheries**
Lamna nasus, porbeagle shark	Pelagic Cold temperate	IUCN: Globally Vulnerable, WNAtlantic Endangered, ENAtlantic Critically endangered CITES: Appendix II	Targeted commercial and recreational, bycatch
Carcharhinus leucas, bull shark	Coastal, estuarine, and freshwater Tropical	IUCN: Globally Near threatened	Targeted commercial and recreational
Sphyrna mokarran, great hammerhead	Coastal-pelagic Tropical	IUCN: Globally Endangered CITES: Appendix II	Targeted commercial and recreational, bycatch
Sphyrna zygaena, smooth hammerhead	Coastal-pelagic Cool temperate to tropical	IUCN: Globally Vulnerable CITES: Appendix II	Targeted commercial, bycatch

References

Au DW, Smith SE, Show C (2009) Shark productivity and reproductive protection, and a comparison with teleosts. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, EA Babcock). Blackwell Publishing Ltd., New Jersey.

Baum JK, Blanchard W (2010) Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research, 102, 229-239.

Baum JK, Kehler DG, Myers RA (2005) Robust estimates of decline for pelagic shark populations in the northwest Atlantic and Gulf of Mexico. Fisheries, 30, 27-29.

Baum JK, Myers RA (2004) Shifting baselines and the decline of pelagic sharks in the Gulf of Mexico. Ecology Letters, 7, 135-145.

Blower D, Pandolfi J, Bruce B, Gomez-Cabrera M, Ovenden J (2012) Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic dispersal events and low effective population sizes. Marine Ecology Progress Series, 455, 229-244.

Calliet GM, Musick JA, Simpfendorfer CA, Stevens JD (2005) Ecology and life history characteristics of Chondrichthyan fish. In: Sharks, Rays, and Chimaeras: The status of the Chondrichthyan fishes (eds. SL Fowler*, et al.*). IUCN, Gland, Switzerland and Cambridge, UK.

Castro JI, Woodley CM, Brudek RL (1999) A preliminary evaluation of the status of shark species. In: FAO Fisheries Technical Paper 380. Food and Agriculture Organization, Rome.

Cortes E (1999) Standardized diet compositions and trophic levels of sharks. ICES Journal of Marine Science, 56, 707-717.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW (2012) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). PLOS One, 7, e29986.

Donoghue PCJ, Purnell MA (2005) Genome duplication, extinction and vertebrate evolution. Trends in Ecology & Evolution, 20, 312-319.

Dudgeon CL, Broderick D, Ovenden JR (2009) IUCN classification zones concord with, but underestimate, the population genetic structure of the zebra shark *Stegostoma fasciatum* in the Indo-West Pacific. Molecular Ecology, 18, 248-261.

Excoffier L, Heckel G (2006) Computer programs for population genetics data analysis: a survival guide. Nature Reviews Genetics, 7, 745-758.

Ferretti F, Worm B, Britten GL, Heithaus MR, Lotze HK (2010) Patterns and ecosystem consequences of shark declines in the ocean. Ecology Letters, 13, 1055-1071.

Gaida IH (1997) Population Structure of the Pacific Angel Shark, *Squatina californica* (Squatiniformes: Squatinidae), around the California Channel Islands. Copeia, 1997, 738-744.

Karl SA, Castro ALF, Garla RC (2012) Population genetics of the nurse shark (*Ginglymostoma cirratum*) in the western Atlantic. Marine Biology.

Keeney D, Heupel M, Hueter R, Heist E (2005) Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. Molecular Ecology, 14, 1911-1923.

Mendonça F, Oliveira C, Gadig OF, Foresti F (2011) Phylogeography and genetic population structure of Caribbean sharpnose shark *Rhizoprionodon porosus*. Reviews in Fish Biology and Fisheries, 21, 799-814.

Mourier J, Planes S (2013) Direct genetic evidence for reproductive philopatry and associated fine-scale migrations in female blacktip reef sharks (*Carcharhinus melanopterus*) in French Polynesia. Molecular Ecology, 22, 201-214.

Musick JA, Burgess G, Cailliet G, Camhi M, Fordham S (2000) Management of sharks and their relatives. Fisheries, 25, 9-13.

Myers RA, Baum JK, Shepherd TD, Powers SP, Peterson CH (2007) Cascading effects of the loss of apex predatory sharks from a coastal ocean. Science, 315, 1846-1850.

Naylor G, Aschliman N (2013) How many species of living sharks, skates and rays are there, and how did they arise over the course of evolution?

Naylor GJP, Caira JN, Jensen K, Rosana KAM, White WT, Last PR (2012) A DNA Sequence–Based Approach To the Identification of Shark and Ray Species and Its

Implications for Global Elasmobranch Diversity and Parasitology. Bulletin of the American Museum of Natural History, 1-262.

Pardini A*, et al.* (2001) Sex-biased dispersal of great white sharks. Nature, 412, 139-140.

Pearse D, Crandall K (2004) Beyond Fst: An analysis of population genetic data for conservation. Conservation Genetics, 5, 585-602.

Shepherd TD, Myers RA (2005) Direct and indirect fishery effects on small coastal elasmobranchs in the northern Gulf of Mexico. Ecology Letters, 8, 1095-1104.

Sodré D, Rodrigues-Filho LF, Souza RF, Rêgo PS, Schneider H, Sampaio I, Vallinoto M (2012) Inclusion of South American samples reveals new population structuring of the blacktip shark (*Carcharhinus limbatus*) in the western Atlantic. Genetics and Molecular Biology, 35, 752-760.

Venkatesh B*, et al.* (2007) Survey Sequencing and Comparative Analysis of the Elephant Shark (*Callorhinchus milii*) Genome, p. e101.

Walker TI (1998) Can shark resources be harvested sustainably? A question revisited with a review of shark fisheries. Mar. Freshwater Res., 49, 553-572.

Whitney NM, Robbins WD, Schultz JK, Bowen BW, Holland KN (2012) Oceanic dispersal in a sedentary reef shark (*Triaenodon obesus*): genetic evidence for extensive connectivity without a pelagic larval stage. Journal of Biogeography, 39, 1144-1156.

CHAPTER 1: The endangered porbeagle shark (*Lamna nasus*): Global population genetic structure, genetic divergence and development of population-level forensic tools

Abstract

Porbeagle (*Lamna nasus*) are an epipelagic shark found in cool temperate waters throughout the Southern Hemisphere and the North Atlantic but are absent from warmer equatorial waters. Commercial fisheries in the North Atlantic target porbeagle for their meat, which is highly valued for human consumption. Due to documented declines in the North Atlantic, porbeagle have been assessed by the IUCN Redlist (International Union for Conservation of Nature) as critically endangered in the eastern North Atlantic and Mediterranean, and endangered in the western North Atlantic. Porbeagle were approved for listing on CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) Appendix II effective September 2014. I investigated porbeagle global genetic structure using the complete mitochondrial control region (mtCR) sequence from individuals (n = 224) throughout the species range. Samples were obtained from three collection areas in the North Atlantic and five collection areas in the Southern Hemisphere. I also sequenced nuclear ITS2 from a subset of these samples and included these mtCR and ITS2 sequences as well as previously published mitochondrial COI sequences in my analyses. I found strong geographic subdivision between the North Atlantic and the Southern Hemisphere, no genetic connectivity between the two hemispheres, no genetic structure within either hemisphere, and among the highest levels of diversity reported for sharks to date. Finally, I expanded testing of a previously reported species ID PCR multiplex and developed a novel PCR multiplex that can determine the hemisphere of origin of porbeagle specimens. The geographic patterns of genetic structure provide a context for regional conservation efforts. The combination of species ID and population ID PCR multiplexes can facilitate catch and trade monitoring of this endangered species.

1. Introduction

Species that have undergone large declines in population size should be evaluated quantitatively to delineate population and stock status, as well as genetic diversity within and among populations. Genetic analyses are a critical component of stock assessment (FAO 2000; Shaklee & Currens 2003; Heist 2005), enabling identification of genetically distinct stocks and populations. Authors have defined populations and stocks with varying criteria and emphasis (Dizon *et al.* 1992; Moritz 1994) although they are generally considered to be at least partially reproductively isolated. Genetically distinct populations typically describe groups of individuals within a species that have distinct evolutionary histories that may infer differing evolutionary potential. A stock can be considered a short-term management unit that typically has a geographical component. Knowledge of genetic population structure and evolutionary history can be critical in managing a species to conserve its genetic diversity as well as the species' ability to adapt to changing environmental conditions (Carvalho & Hauser 1994; Hilborn *et al.* 2003).

The annual catches of most pelagic fish meet or exceed the estimated maximum sustainable yields and sharks are no exception (Baum & Myers 2004; Baum *et al.* 2005; Myers & Worm 2005; Worm *et al.* 2005; Baum & Blanchard 2010). Many populations of elasmobranches decline more quickly and recover more slowly than marine teleosts in response to overfishing, because they are typically long-lived, mature at a late age, and have long gestations and low fecundities (Cortes 2002; Garcia *et al.* 2008; Au *et al.* 2009; Dulvy & Forrest 2010).

The porbeagle shark (*Lamna nasus*) is found in cold temperate waters throughout the Southern Hemisphere and in the North Atlantic; its sister species the salmon shark (*Lamna ditropis*) is present in the North Pacific. The porbeagle is even less fecund than many other shark species with a mean litter size of 3.7-4 pups per year (Francis *et al.* 2008), thus it is potentially highly sensitive to overexploitation. Unlike that of most sharks, porbeagle meat is highly valued for human consumption and is an important component of commercial fisheries. The species has been heavily fished in targeted commercial fisheries in the North Atlantic and Mediterranean. The porbeagle is not a targeted species in the Southern Hemisphere, but it comprises a significant proportion of

bycatch in other commercial fisheries (Francis *et al.* 2008). Porbeagle abundance has declined to approximately 10-20% of its virgin biomass in the western North Atlantic, with estimated declines to below maximum sustainable yield in the rest of its range (DFO 2005; ICCAT 2009). Indeed, due to historic overfishing the porbeagle is assessed as endangered in the western North Atlantic and critically endangered in the eastern North Atlantic and in the Mediterranean by the IUCN Red List (Stevens *et al.* 2009). A CITES Appendix II listing will be effective September 2014. The current lack of international management highlights the need for genetic evaluation of the species and regional harvest information that may be inferred from such genetic data.

Due to the economic importance of the porbeagle and the widely recognized collapse of fisheries targeting it in the North Atlantic (Stevens *et al.* 2000; Campana *et al.* 2008), the natural history of this shark is of interest to inform fisheries management and conservation efforts. The species is endothermic, prefers water temperatures less than 18^0C, and inhabits the higher latitudes of the North Atlantic and the Southern Hemisphere. Porbeagle have not been documented in equatorial waters (Francis & Stevens 2000; Campana & Joyce 2004; Last & Stevens 2009). This anti-equatorial distribution suggests reproductive isolation between North Atlantic and Southern Hemisphere porbeagle and the hypothesis of genetically divergent northern and southern populations. Supporting this hypothesis are differences in life history characteristics between northern and southern porbeagle, with reported variation in growth rates, size at maturity, maximum size, and longevity (Francis & Stevens 2000; Jensen *et al.* 2002; Francis & Duffy 2005; Francis *et al.* 2007; Francis *et al.* 2008). The potential reproductive isolation combined with demographic differences have raised the question of whether porbeagle comprise more than one species, but this issue has not been explored genetically.

Identification of sharks present in fisheries and markets is difficult because morphological characters vary only slightly between many species. This is compounded by the common practice of finning in which the fins are retained and the remainder of the carcass with its potentially distinguishing morphological features is discarded at sea to maximize storage space on board the fishing vessels. These two factors make monitoring the catch and trade of sharks at the species level extremely difficult. Molecular genetic

techniques have been developed to identify porbeagle and other shark species when morphologic characters are insufficient or unavailable. Shivji et al. (2002) developed a PCR multiplex assay based on the nuclear ribosomal internal transcribed spacer 2 (ITS2) locus which differentiates porbeagle from 5 other pelagic species commonly found in global fisheries, but were only able to test the accuracy of this assay using 17 porbeagle from the western North Atlantic. More recently, Wong et al. (2009) developed a character-based diagnostic system for shark species identification, including porbeagle, utilizing the DNA barcode portion of the mitochondrial cytochrome oxidase I (COI) gene. Although the delineation of genetic population structure could provide important information for the conservation and management of this endangered species, such analyses were beyond the scope of these two previous studies.

Here I assess the population genetic structure of the porbeagle and its applicability and utility for conservation purposes. I present the results of analyses of DNA sequence variation in the mitochondrial control region (mtCR) of porbeagle sampled throughout the species' global range. I specifically evaluate (1) the overall genetic structure of the species and the remaining levels of genetic diversity, (2) the amount of genetic differentiation within and between the hemispheres, (3) the demographic history and taxonomic status of the species, and (4) whether genetic data can be used to identify the geographic origin of porbeagle in trade. Finally, I expanded testing of a PCR multiplex used to identify porbeagle to the species level (Shivji *et al.* 2002).

2. Materials and Methods

2.1. Samples, DNA extraction, amplification and sequencing

A total of 224 specimens were obtained from other researchers, fisheries observer programs and artisanal and recreational fishers from three areas in the North Atlantic and five areas throughout the Southern Hemisphere (Figure 1). The collection areas and their geographic abbreviations used herein are 1) North Atlantic (NAtl): western North Atlantic (WNA), Denmark (DK), and United Kingdom (UK); and 2) Southern Hemisphere (SH): Chile (CH), Falkland Islands (FKI), South Africa (SAfr), New Zealand (NZ), and Tasmania (Tas). Due to the small sample sizes from the UK, Falkland

Islands and Tasmanian collection areas, data from those collection areas were not used in statistical analyses or were combined with data from nearby collection areas.

Genomic DNA was isolated from tissue samples using the QIAGEN DNeasy Tissue kit (QIAGEN, Inc., Valencia, CA). The entire mtCR plus some flanking DNA was amplified by primers LnasCRF6 and DasR2 (primer details in Supplementary Information, Table S1). PCR reactions were performed in 50 μl volumes containing 1μl of unquantified genomic DNA, 200 μM of each dNTP, 12.5 pM each primer, 1 U HotStar *Taq*^{TM} DNA polymerase (QIAGEN, Inc.) and 5μl 10x HotStar *Taq*^{TM} reaction buffer. PCR cycling conditions were 95ºC for 15 minutes, 35 cycles of 94ºC for 1 minute, 50ºC for 1 minute and 72ºC for 2 minutes, with a final extension of 5 minutes at 72ºC. In each set of PCR amplifications, a negative control with no genomic DNA was included to check for contamination.

Amplified products were purified using the QIAquick PCR Purification Kit (QIAGEN, Inc.) prior to direct cycle sequencing with BigDye 3.1 Terminator chemistry (Applied Biosystems, Inc., Foster City, CA) on both strands using primers CRF6, DasR2, LnasCRF8 and LnasCRR5. Sequencing reactions were purified using Dyex 2.0 Spin Kit (QIAGEN, Inc.) and sequenced on an AB3130 genetic analyzer (Applied Biosystems, Inc.). Sequences were aligned with GENEIOUS version 4.04 (Drummond *et al.* 2008) and alignments were checked and finalized by eye. Novel sequences were submitted to GenBank (accession numbers --- to ---).

2.2. Genetic diversity and population structure

The ARLEQUIN software version 3.01 (Excoffier *et al.* 2005) was used to calculate molecular diversity indices such as the number of haplotypes (*nh*), haplotype diversity (*h*) and nucleotide diversity (π). For all statistical analyses with multiple tests, I evaluated significance levels ($P < 0.05$) after sequential Bonferroni correction.

I investigated genetic structure within a geographic context using three different approaches. First, the amount of genetic differentiation within and among populations was examined by an analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) with the Tamura & Nei (Tamura & Nei 1993; Bowen *et al.* 2005) distance as implemented in ARLEQUIN. AMOVA calculates analogs to Wright's F-statistics (Wright

1951, 1965), designated Φ_{ST}, based on the allelic content of haplotypes and haplotype frequencies. I initially calculated pairwise Φ_{ST} values among the five collection areas with n > 10 samples. I pooled samples from neighboring sampling areas that had non-significant pairwise Φ_{ST} values ($P > 0.05$) into putative populations for subsequent population level or statistical analyses. I evaluated the amount of genetic differentiation within and among these populations through hierarchical AMOVA. The significance of the covariance components was assessed using 1 000 non-parametric permutations. Secondly, I estimated genetic differentiation between collection areas using Hudson's nearest neighbor statistic (Snn) (Hudson 2000) as calculated in DNASP version 5.0 (Librado & Rozas 2009) with 1 000 permutations. Finally, I conducted SAMOVA analyses of pairwise differentiation (Dupanloup *et al.* 2002) with a range of K (2 - 5) and 1000 permutations.

Haplotypes were identified using DNACOLLAPSER version 1.0 (Fredsted 2006). Genealogies were estimated by unrooted statistical parsimony networks constructed using TCS version 1.21 (Clement *et al.* 2000) with the connection limit initially set at 95%. In cases where multiple sub networks were present, the connection limit was scaled down from 95% to 90% to determine whether the networks would join at a lower limit. Ambiguities in the networks were resolved using criteria based on coalescent theory (Crandall & Templeton 1993; Pfenninger & Posada 2002); alternate network connections are not shown.

2.3. Genetic Distance and Mutation Rate Estimation

MEGA version 3.1 (Kumar *et al.* 2004) was used to calculate the pairwise genetic distances (d) and within group means (π) within and among groups under the Tamura Nei evolution model. The variance within groups was corrected by d corr = $d - ((\pi_1 - \pi_2)/2)$ and the corrected distance was divided by the time since divergence to estimate a mutation rate μ. Mutation rates were estimated based on divergence of the mtCR sequence among several extant lamnid, carcharhinid and sphyrnid species and comparison with (1) the time of first appearance in the fossil record or (2) comparisons of inferred phylogeny and the fossil record (Martin *et al.* 2002).

2.4. Demographic History

A variety of approaches were used to evaluate population demographics. Estimates of mutation rates and generation times are required to translate parameter estimates into demographic terms. I used a range of mutation rates both previously published and calculated herein for my analyses. It has been reported that 50% of Northwest Atlantic female porbeagle are mature at age 13 (Jensen *et al.* 2002; Natanson *et al.* 2002) and that Southern Hemisphere females mature at 15 – 18 years (Francis & Duffy 2005; Francis *et al.* 2007). I used 13 years as the generation time, a generation time of 18 years decreased divergence time and effective population size estimates (data not shown).

I first calculated demographic summary statistics in ARLEQUIN. Tajima's *D* statistic (Aris-Brosou & Excoffier 1996; Tajima 1996) tests the hypothesis of selective neutrality and population equilibrium; however significant *D* values can also result from factors such as heterogeneity of mutation rates. Fu's F_S (Fu 1997) evaluates neutrality of mutations, and is very sensitive to population expansion which leads to large negative values of F_S. The mismatch distribution analysis (Schneider & Excoffier 1999) in ARLEQUIN computes the number of differences between pairs of haplotypes and uses a non-linear least-square approach to estimate parameters of a sudden demographic (Rogers & Harpending 1992) and geographic (Ray *et al.* 2003) expansion. The sum of squares deviations (*SSD*) between the observed and expected mismatch distribution tests the validity of the expansion model. Harpending's raggedness index (*Hri*) (Harpending 1994) tests the fit of the mismatch distribution and provides indications of population expansion, with larger values of *Hri* typical of multimodal distributions found in stationary populations and smaller, non-significant values for unimodal or smoother distributions typical of expanding populations. Tau (τ) is calculated in the mismatch analysis and is a relative measure of the time in generations since population expansion ($\tau = 2\mu t$). I also report time since expansion in years (Time) calculated by $t = \tau/2\mu$ where μ is the mutation rate per site per generation. For the geographic expansion model, *m* is the migration rate between the sampled deme and a population of infinite size after T generations, $M = 2N_e m$, and N_e is the effective population size.

MDIV (Nielsen & Wakeley 2001), a Markov chain Monte Carlo method, was used to estimate Θ ($\Theta = N_e\mu$), migration rates ($m = \Theta M/2$), divergence time ($t=2T\ N_e$), and

estimated time to most recent common ancestor (TMRCA) between the North Atlantic and Southern Hemisphere populations of *L. nasus*. Isolation is inferred when M~ 0 and T is unimodal, migration is inferred when M > 0. I used the HKY model of evolution because, according to the software authors, it is a more accurate model of DNA evolution although it is computationally slower than the other option available in the program. Chain length varied from 2 x 10^6 to 1 x 10^7 cycles and burn-in time was set to 10% of the total length of cycles. For the initial data collection runs, Mmax was set at 100 and Tmax was set at 20 to determine appropriate upper bounds. For each population or collection area comparison, the data set was analyzed at least three times with different random seeds to evaluate convergence. The value with the highest likelihood was selected as the best estimate of each output parameter.

The Bayesian implementation in MIGRATE (Beerli & Felsenstein 2001; Beerli 2006), a coalescent-based Markov chain Monte Carlo method, was also used to estimate Θ (Θ = $N_e\mu$) for both populations and migration rates (m = ΘM/2) between hemispheres. Pairwise comparisons within both populations were performed for collection areas with n ≥ 30 samples. Each data set was analyzed at least three times with different random seeds while the final input prior values and all other inputs were held constant to evaluate convergence. For each of the output parameters, the means from each of the final runs were averaged to obtain the best estimate.

I inferred changes in effective population size over time and estimated current effective population sizes through Bayesian skyline plots as implemented in BEAST v1.7 (Drummond *et al.* 2012). I used the HKY+I+G model because it most closely approximated the best fit model of DNA evolution estimated using the AICc in MODELTEST v3.7 (Posada & Crandall 1998). I ran three separate analyses for each hemisphere using a linear skyline model with 5 groups, genealogies and parameters sampled every 1 000 iterations, and automatic optimization of operators. For the North Atlantic I ran the analyses for 2.3*10^8 iterations with a strict molecular clock. For the Southern Hemisphere I ran the analyses for 1.5 x 10^8 iterations with a relaxed lognormal molecular clock. Convergence was assumed when the ESS for each parameter in each replicate run exceeded 200. I generated skyline plots in TRACER v1.5 (Drummond & Rambaut 2007).

2.5. Intra- and inter-specific divergences

I evaluated the implications of the inter-hemisphere genetic differentiation by comparing intra- and inter-specific genetic divergences. In addition to my mtCR sequence data, I analyzed two loci that have been used for porbeagle species identification in other studies, mitochondrial COI (Wong *et al.* 2009) and nuclear ITS2 (Shivji *et al.* 2002). To evaluate the amount of inter-hemispheric variation in these loci, I re-analyzed the 80 porbeagle COI sequences (30 North Atlantic and 50 Southern Hemisphere) from Wong et al. (2009) and sequenced the entire nuclear ITS2 from 21 porbeagle individuals (12 North Atlantic and 9 Southern Hemisphere). The complete ITS2 was amplified using primers Fish58SF and Fish28SR (Shivji *et al.* 2002). Amplification conditions were the same as in section 2.1, except the annealing temperature used was 65ºC. Double stranded sequence data was generated using primers Fish58SF, Fish28SR, Lnas556F, Lnas827F, Lnas485R, and Slmn739R (primer details in Supplementary Information, Table S1), following the sequencing reaction conditions in section 2.1. Statistical parsimony networks were constructed from the 80 porbeagle COI sequences (Wong *et al.* 2009) and the 19 porbeagle ITS2 sequences as described in section 2.1.

I then calculated comparative inter-specific genetic distances using mtCR, COI and ITS2 sequences from other lamnid, carcharhinid and sphyrnid shark species. All COI sequences were from Wong et al. (2009), the *Carcharodon carcharias* mtCR sequences were obtained from GenBank (accession numbers AU022396, AY026207, and AY026212), and all other mtCR and ITS2 sequences were generated as described using primers listed in Supplementary Information, Table S1. All novel sequences were deposited in GenBank (accession number xxx – xxx).

2.6. Market Sample Origin and Species ID

An additional 14 samples were obtained from the Hong Kong fin market and 16 from fin traders in South Africa. The mtCR from these market samples was sequenced and aligned with porbeagle sequences from known collection areas in Geneious as described above. A statistical parsimony network was created as described above but with the connection limit set at 93% to join the northern and southern population subnetworks.

A total of 17 western North Atlantic porbeagle reference specimens were previously tested using an ITS2 species ID multiplex PCR (Shivji *et al.* 2002). I tested an

additional 83 North Atlantic and 100 Southern Hemisphere porbeagle samples following those conditions.

To design hemisphere specific primers I used the mtCR haplotype alignment. Several PCR primers putatively specific for each population were designed incorporating nucleotide sequence differences in the mtCR. To test and optimize each primer, I used triplex PCR reactions that included porbeagle forward (Lnas125F) and reverse (Lnas865R) primers together with one putatively hemisphere specific primer, sensu Shivji et al. (2002). My expectation was that the three primer test for Northern Hemisphere porbeagle would yield two amplicons when used to amplify Northern Hemisphere porbeagle genomic DNA: (1) an 800 bp positive control amplicon generated from the porbeagle forward and reverse primer and (2) a smaller amplicon diagnostic for the Northern Hemisphere porbeagle. When this three primer test for Northern Hemisphere porbeagle was used to amplify Southern Hemisphere porbeagle genomic DNA it would yield only the positive-control amplicon. My expectations for the three primer test for Southern Hemisphere porbeagle followed similar logic.

After the Northern (Lnas556F-3) and Southern (Lnas293F-2) Hemisphere specific primers were optimized, they were combined into a single quadraplex PCR reaction with porbeagle forward (Lnas125F) and reverse (Lnas865R) primers. When this multiplex PCR was used to amplify porbeagle genomic DNA, the expectation was that an 800 bp positive control amplicon generated from the porbeagle forward and reverse primers would always be present, with one of two smaller amplicons present depending on whether the genomic DNA was from a Northern Hemisphere (400 bp) or Southern Hemisphere (650 bp) porbeagle. Reaction conditions for the species specific and hemisphere specific amplifications were as described in Shivji et al. (2002).

3. Results

3.1. Sequence characteristics, genetic diversity and population genetic structure

The porbeagle mtCR sequence is 1 063 nucleotides long. The complete mtCR from 224 individuals identified a total of 69 variable sites including 51 parsimony informative sites and 7 variable sites that are geographically fixed between the North Atlantic and Southern Hemisphere. These constituted 100 haplotypes that segregated completely by

geography with 54 haplotypes in the North Atlantic and 46 in the Southern Hemisphere (Supplementary Information, Table S2). Patterns of haplotype frequency were different in the hemispheres. In the North Atlantic, only one haplotype was present in more than 3 individuals. Two haplotypes were most common in the Southern Hemisphere, each present in more than 40 individuals, and two others were present in more than 5 individuals.

The COI sequences (Wong *et al.* 2009) I re-analyzed were 645 nucleotides in length. Two variable sites were geographically fixed between hemispheres. The 80 sequences collapsed to 22 haplotypes for a haplotype diversity of 0.846. No haplotypes were shared between hemispheres. The ITS2 sequences were 1 144 nucleotides and had 12 variable sites. Ten of these variable sites displayed differential polymorphism in that all individuals in one hemisphere were homozygotes with a single peak at those positions while some individuals (two sites) or all individuals (eight sites) in the other hemisphere were heterozygous with two peaks of equal intensity. The two sites with differing levels of polymorphism were excluded from analysis, leaving ten variable sites that segregated by hemisphere. The ITS2 sequences collapsed to 2 genotypes that were not shared between hemispheres.

The mtCR was more variable than either COI or ITS2, thus mtCR was used for detailed population genetic analyses. I report population and demographic statistics from ARLEQUIN analyses (Table 1) by collection area, population, and overall global rates. Overall haplotype and nucleotide diversity were very high in porbeagles at 0.929 and 0.0133, respectively.

Pairwise Φ_{ST}s (Table 2a) were all large and significant ($\Phi_{ST} > 0.75$, $P < 0.00001$) in comparisons of North Atlantic to Southern Hemisphere collection areas before and after sequential Bonferroni correction. Pairwise Φ_{ST}s between collection areas within hemispheres were small and non-significant ($\Phi_{ST} < 0.00$, $P > 0.40$) so sequences from collection areas within hemispheres were grouped together. The AMOVA analyses revealed significant genetic structure between the North Atlantic and Southern Hemisphere ($\Phi_{ST} = 0.83$, $P < 0.00001$), so hereafter they are treated as two genetically distinct populations. Heirarchical AMOVA (Table 2b) revealed that 82.7% of the variation existed between the populations, 17.6% was within the collection areas, and

there was essentially no variation among collection areas within populations (-0.36%). Similarly, inter-hemispheric pairwise nearest neighbor statistics (Table 2a) were all equal to 1 and statistically significant ($P < 0.05$) before and after correction indicating strong differentiation between the populations, while intra-hemispheric pairwise Snn's were all approximately 0.5 and non-significant ($P > 0.05$), indicating no differentiation among collection areas within either population. SAMOVA analyses (Table 2c) indicate that between group variance was greatest when K = 2, which grouped the North Atlantic collection areas together and separate from a group comprising the Southern Hemisphere collection areas ($\Phi_{CT} = 0.82$, $P = 0.0062$).

Statistical parsimony networks (Figure 2) are consistent with the differentiation between and the lack of genetic structure within populations that was evident in the AMOVA, SAMOVA and nearest neighbor analyses. The majority of haplotypes obtained from more than one individual were present in individuals from different collection areas, no genetic structure was evident within either population, and the two populations shared no haplotypes. The North Atlantic and Southern Hemisphere mtCR clades did not join at the 95% connection limit so the limit was decreased until the clades were joined at the 93% connection limit by 15 mutation steps. The Southern Hemisphere clade was dominated by 2 haplotypes with a star-like network radiating from both. Both clades have a number of alternate connections that are not shown. Statistical parsimony networks of the COI and ITS2 sequences (not shown) produced similar topologies with differentiation between and lack of spatial genetic structure within the hemispheres. The North Atlantic and Southern Hemisphere COI and ITS2 clades do join at the 95% connection limit.

3.2. Genetic Distance and Mutation Rate Estimates

Genetic distances (Supplementary Information, Table S3) between several species of the families Carcharhinidae, Sphyrnidae and Lamnidae and the mutation rates (Supplementary Information, Table S4) based on mtCR distances are presented in Supplementary Information. The mtCR genetic distance (Supplementary Information, Table S3a) between the two porbeagle populations was 0.020 while genetic distance between pairs of species ranged from 0.051 between porbeagle and salmon sharks (*Lamna ditropis*), 0.074 between the shortfin (*Isurus oxyrhinchus*) and longfin (*Isurus*

paucus) makos, to 0.169 between great white (*C. carcharias*) and salmon sharks. Genetic distance ranged from 0.036 to 0.064 between pairs of Carcharhinid species and from 0.138 to 0.197 between pairs of Sphyrnids. The COI genetic distance between the shortfin and longfin makos is 0.143, between porbeagle and salmon sharks is 0.051, and between the two porbeagle populations is 0.007. Genetic distances based on COI and ITS2 sequences are presented in Supplementary Information Tables S3b and S3c, respectively. My calculated mutation rates include 0.64% per million years based on the divergence of Sphyrnid and Carcharhinid sharks 38 mya (Maisey 1984; Cappetta 1987); 0.38% per million years for comparisons of *C. leucas* specimens from the eastern Pacific and the Gulf of Mexico; and an average of 0.23% per million years for comparisons between the extant Lamnid shark species.

3.3. Population demographics

Fu's F_S was negative and significant for all collection areas with n > 30 (Table 3) and for the two populations separately and combined. The mismatch distributions (*SSD*) and *Hri* did not differ significantly from that expected under population expansion ($P > 0.05$) and both *SSD* and *Hri* values for the North Atlantic were smaller than values for the Southern Hemisphere. The difference between Θ_0 and Θ_1 was large for all collection areas and both populations. The τ-values were similar at collection areas within populations, and τ-values in the North Atlantic were almost twice those in the Southern Hemisphere.

Results of the MDIV and MIGRATE analyses indicate that divergence between the two populations occurred during the Pleistocene and the subsequent migration rate between populations has been less than 1 individual per generation (Table 4). For adjacent pairs of collection areas within a population, estimated divergence times and migration rates greater than 10 individuals per generation were consistent with no population structure within the populations as detected by the AMOVA analyses. However, due to the multimodality of migration rates between collection areas I have not inferred directionality of migration events within populations. Estimates of Θ were generally concordant across the methods.

Bayesian skyline plots (Figure 3) also revealed expansion in both populations during the Pleistocene. Expansion started during the middle Pleistocene in the North Atlantic and may have peaked in the late Pleistocene or Holocene. Expansion started

more recently in the Southern Hemisphere, sometime during the late Pleistocene, and appears to be ongoing through the Holocene.

3.4. Species ID and population of origin

The mtCR sequences from 16 fins from the South Africa fin traders clustered within the Southern Hemisphere clade while the mtCR sequences from 14 fins from the Hong Kong market clustered within the North Atlantic clade (Figure 2). Representative results of the species (ITS2) and hemisphere (mtCR) multiplexes are shown in Figure 4. All porbeagle specimens from both hemispheres were correctly identified as porbeagle with no misidentifications (Figure 4a). The mtCR hemisphere multiplex correctly assigned porbeagle specimens to their hemisphere of origin (Figure 4b).

4. Discussion

Documented declines in porbeagle abundance highlight the need for evaluation of the genetic population structure, the genetic diversity, and the taxonomic status of this species. This study provides a detailed genetic analysis of this commercially important, endangered shark species with solid sample coverage throughout most of its geographic range. I analyzed the complete mtCR of 224 porbeagle from 8 collection areas and an additional 30 samples from fin trade markets. I also evaluated mitochondrial COI (Wong *et al.* 2009) and nuclear ITS2 sequences from a subset of these sampled individuals.

4.1. Population structure and migration rates

The porbeagle shark exhibits strong genetic subdivision between the North Atlantic and Southern Hemisphere populations and no genetic structure within populations. Levels of genetic diversity within both populations and the species overall are among the highest of any shark when compared to a recent summary of reported elasmobranch mtCR studies (Karl *et al.* 2011). My results are consistent with no gene flow between the two populations while estimated migration rates among collection areas within populations are greater than 10 individuals per generation.

In contrast, long-term conventional tagging data suggests that there are separate stocks in the eastern and western North Atlantic with no appreciable mixing between them (Stevens 1990; Campana *et al.* 2001; Kohler *et al.* 2002). Only one movement of

an individual between the eastern and western North Atlantic has been observed in conventional tag-recapture studies (Francis *et al.* 2008). Similarly, recent satellite tagging studies of 28 porbeagle in the North Atlantic reported movements of similar distances with all sharks remaining in their general capture area (Pade *et al.* 2009; Campana *et al.* 2010; Saunders *et al.* 2010). Although the eastern and western North Atlantic have been considered to be two separate stocks for fisheries management purposes, I find no genetic evidence for differentiation between the eastern and western North Atlantic porbeagle. The lack of trans-Atlantic movement evidenced by tagging studies is not necessarily inconsistent with my finding of no spatial genetic partitioning within the North Atlantic. As noted by Boustany et al. (2008), tagging studies cannot determine with certainty that genetic isolation exists because untagged fish may move between areas, thus genetic studies can be more informative in delineating population structure. On average, one gene copy per generation will prevent substantial divergence at a locus (Wright 1931). My data suggest approximately 30 - 150 migrants per generation between the eastern and western North Atlantic or roughly 2 - 12 migrants per year assuming a 13-year generation time. These estimates are large enough to facilitate gene flow between the two stocks and potentially confound any expected genetic structure that might otherwise exist, yet small enough to have escaped detection by tagging studies because of the limited number of recapture events. There are no reported conventional or satellite tagging studies for porbeagle in the Southern Hemisphere.

4.2. Demographic history and taxonomic relationships

Our data indicate an older coalescence in the North Atlantic with a northern population expansion in the early to middle Pleistocene followed by colonization of the Southern Hemisphere during the first Pleistocene glacial period 0.8 - 1.3 mya. Additional migration events between hemispheres may have occurred during more recent glacial periods when the equatorial Atlantic waters were cooler. The likely cause of the starlike phylogeny in the Southern Hemisphere is a dramatic population expansion over the last 90,000 - 350,000 years as seen in the Bayesian skyline plots. Thus my data are consistent with a hypothesized porbeagle origin in the North Atlantic in the late Cenozoic, followed by colonization of the Southern Hemisphere during a Pleistocene glacial period (Reif & Saure 1987).

Interspecific or interpopulation genetic distances were calculated for several shark species and compared to (1) their first appearance in the fossil record, or (2) comparisons of inferred phylogeny and the fossil record (Martin *et al.* 2002). My estimated mutation rates are comparable to recently published rates based on comparisons of shark populations or species separated by the Isthmus of Panama (Duncan *et al.* 2006; Keeney & Heist 2006; Schultz *et al.* 2008; Nance *et al.* 2011) and may suggest mutation rate differences among elasmobranch families.

Our genetic distance calculations corroborate traditional systematics in that both pairs of *Isurus* and *Lamna* species appear to be more closely related to each other than to members of other genera within the Lamniformes. Genetic distances between the porbeagle populations at all three loci are approximately half that between porbeagle and their congener salmon sharks (*L. ditropis*) and less than half that between other Lamnids (Table S3). My observed COI genetic distance of 0.7% between northern and southern porbeagle is considerably lower than the proposed 2 - 3% threshold for species delineation in COI barcoding (Hebert *et al.* 2003). Inter-specific divergence is typically 1 – 2 orders of magnitude greater than intra-specific variation at mitochondrial loci (Hebert *et al.* 2003; Greig *et al.* 2005). In both mitochondrial loci the inter-population divergence is less than 1 order of magnitude greater than the intra-population diversity. Thus my genetic distances indicate that the divergence between the two porbeagle populations is not great enough to infer a speciation event. In contrast, a recent literature survey found that in general, no matter what locus was used, DNA sequences from a single species typically form a single statistical parsimony network at the 95% connection limit while sequences from different species separate into different 95% networks (Hart & Sunday 2007). The COI and ITS2 porbeagle sequences both comprise single networks at the 95% connection limit, whereas networks from all three loci do not join porbeagle and salmon sharks at even the 90% connection limit. My observation that the porbeagle mtCR northern and southern sub-networks join at the 93% but not at the 95% connection limit may indicate that the two populations are separate species. However, Hart and Sunday (2007) also found that the mean rate of concordance between the number of taxa and the number of 95% sub-networks was considerably higher in studies that utilized COI than in all other studies, thus I place more weight on my single 95% COI network

than on the mtCR having two subnetworks at the 95% level. After considering all of these factors I believe that my data indicates that North Atlantic and Southern Hemisphere porbeagles are two genetically distinct populations of the same species, but recommend studies using additional nuclear loci to further clarify this point.

4.3. Species identification and population of origin

Shivji et al. (2002) reported a multiplex PCR reaction that can distinguish between porbeagle and five other sharks commonly encountered in pelagic fisheries and the global fin trade. At that time only 17 porbeagle samples from the western North Atlantic were available. I tested this species ID multiplex on a total of 100 samples from the North Atlantic and 100 from the Southern Hemisphere. All were correctly identified unambiguously as porbeagle. The sequence variation observed in both ITS2 and COI does not affect the use of these loci in species identification. The mtCR of porbeagle sharks can be used to determine whether a specimen originated in the North Atlantic or Southern Hemisphere, but does not provide resolution at a finer scale. I developed a novel hemisphere diagnostic PCR multiplex based on the mtCR sequence data, tested it on the same 200 samples, and found that it correctly determines the hemisphere of origin. The combination of the species and hemisphere identification multiplexes enables rapid identification of porbeagle in trade and determination of the hemispheric origin.

5. Conclusions and implications for management and conservation

The results of these analyses indicate that the North Atlantic and Southern Hemisphere represent two genetically distinct populations of porbeagle, although there is not enough genetic divergence to suggest that they are different species. The estimated intrapopulation migration rates are low but sufficient to promote gene flow and a corresponding lack of genetic differentiation within both populations. However, regional recovery based on recruitment from nearby locales may not occur because a much larger number of migrants per generation is required to replenish depleted stocks than to establish gene flow (Waples 1998). I propose that the Northern and Southern Hemisphere be managed as two separate, genetically distinct populations. Although no genetic differentiation was found between the eastern and western North Atlantic porbeagle, recruitment between these areas is potentially low and they should be

considered two stocks for fisheries management purposes. Porbeagle in the Southern Hemisphere appear to comprise a single panmictic population, although migration rates among collection areas appear to be too low to facilitate meaningful recruitment. The available data are insufficient to define the appropriate number or location of management stocks in the Southern Hemisphere.

The genetic differentiation between populations makes it possible to determine the hemisphere of origin of porbeagle in trade. The ITS2 and mtCR multiplex PCR tests are rapid and inexpensive, and can facilitate catch and trade monitoring as well as provide insight into the geographic utilization of the species. Although significant declines in porbeagle population size are well documented, there does not appear to be a corresponding decline in genetic diversity. This is likely due to the relatively recent increase in fishery efforts, equivalent to only 2 – 3 generations ago, which may not be enough time for an impact to be observed in the genetic signature. Additionally, the species has a relatively long lifespan, and with overlapping generations many of the sampled animals may have lived through the population decline. The remaining high level of genetic diversity bodes well for the future recovery of the species if aggressive management practices are implemented quickly.

Acknowledgements

I thank the following individuals for porbeagle samples used in this study: Debra Abercrombie, Shelley Clarke, Ken Collins, Malcolm Francis, Denise Headon, Catherine Jones, Ben King, Sebastian Hernandez Munoz, Lisa Natanson, Les Noble, Nick Pade, and Richard Pierce. For their laboratory and analytical support I thank Demian Chapman, Jennifer Hester, Lara Murphy, and Vince Richards. Financial support for this project was received from Florida Sea Grant.

References

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Moleclar Biology and Evolution, 13, 494-504.

Au DW, Smith SE, Show C (2009) Shark productivity and reproductive protection, and a comparison with teleosts. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, EA Babcock). Blackwell Publishing Ltd., New Jersey.

Baum JK, Blanchard W (2010) Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research, 102, 229-239.

Baum JK, Kehler DG, Myers RA (2005) Robust estimates of decline for pelagic shark populations in the northwest Atlantic and Gulf of Mexico. Fisheries, 30, 27-29.

Baum JK, Myers RA (2004) Shifting baselines and the decline of pelagic sharks in the Gulf of Mexico. Ecology Letters, 7, 135-145.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics, 22, 341-345.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations using a coalescent approach. PNAS, 98, 9157-9160.

Boustany AM, Reeb CA, Block BA (2008) Mitochondrial DNA and electronic tracking reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). Marine Biology, 156, 13-24.

Bowen BW, Bass AL, Soares L, Toonen RJ (2005) Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). Molecular Ecology.

Campana S, Marks L, Harley S (2001) Analytical assessment of the porbeagle shark (*Lamna nasus*) population in the northwest Atlantic, with estimates of long-term sustainable yield. In: Fisheries and Oceans Science, Canada, pp. 1-17.

Campana SE, Joyce W, Fowler M (2010) Subtropical pupping ground for a cold-water shark. Can J Fish Aquat Sci, 67, 769-773.

Campana SE*, et al.* (2008) The rise and fall (again) of the porbeagle shark population in the Northwest Atlantic. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, EA Babcock), pp. 445-461. Blackwell Publishing, Oxford, UK.

Campana SE, Joyce WN (2004) Temperature and depth associations of porbeagle shark (*Lamna nasus*) in the northwest Atlantic. Fisheries Oceanography, 13, 52-64.

Cappetta H (1987) Chondrichthyes II: mesozoic and cenozoic elasmobranchii Gustav Fischer Verlag, Stuttgart, Germany.

Carvalho GR, Hauser L (1994) Molecular genetics and the stock concept in fisheries. Reviews in Fish Biology and Fisheries, 4, 326-350.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9, 1657-1659.

Cortes E (2002) Incorporating Uncertainty into Demographic Modeling: Application to Shark Populations and Their Conservation. Conservation Biology, 16, 1048-1062.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics, 134, 959-969.

DFO (2005) Stock Assessment report on NAFO subareas 3 - 6 Porbeagle shark (ed. DoFa Oceans). Canadian Science Advisory Secretariat.

Dizon AE, Lockyer C, Perrin WF, DeMaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. Conservation Biology, 6, 24-36.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7, 214.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2008) Geneious v4.0.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969-1973.

Dulvy ND, Forrest RE (2010) Life histories, population dynamics, and extinction risks in chondrichthyans. In: Sharks and their relatives II: Biodiversity, adaptive physiology and conservation (eds. JC Carrier, JA Musick, MR Heithaus), pp. 635-676. CRC Press, Boca Raton, FL.

Duncan K, Martin A, Bowen B, DeCouet H (2006) Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*). Molecular Ecology, 15, 2239-2251.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Molecular Ecology, 11, 2571-2581.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics, 2005, 47-50.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131, 479-491.

FAO (2000) Fisheries Management : 1. Conservation and management of sharks. FAO technical guidelines for responsible fisheries 4 (Suppl 1). FAO (United Nations Food and Agriculture Organization) Rome.

Francis MP, Campana SE, Jones CM (2007) Age under-estimation in New Zealand porbeagle sharks (*Lamna nasus*): is there an upper limit to ages that can be determined from shark vertebrae? Marine and Freshwater Research, 58, 10-23.

Francis MP, Duffy C (2005) Length at maturity in three pelagic sharks (*Lamna nasus*, *Isurus oxyrinchus* and *Prionace glauca*) from New Zealand. Fishery Bulletin, 108, 489-500.

Francis MP, Natanson LJ, Campana SE (2008) The biology and ecology of the porbeagle shark, *Lamna nasus*. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, E Babcock), pp. 105-113. Blackwell Publishing, Oxford UK.

Francis MP, Stevens JD (2000) Reproduction, embryonic development, and growth of the porbeagle shark, *Lamna nasus*, in the southwest Pacific Ocean. Fisheries Bulletin, 98, 41-63.

Fredsted PV (2006) FaBox - an online fasta sequence toolbox, http://www.birc.au.dk/fabox.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147, 919-925.

Garcia VB, Lucifora LO, Myers RA (2008) The importance of habitat and life history to extinction risk in sharks, skates, rays and chimaeras. Proceedings of the Royal Society B, 275, 83-89.

Greig TW, Moore MK, Woodley CM, Quattro JM (2005) Mitochondrial gene sequences useful for species identification of western North Atlantic Ocean sharks. Fisheries Bulletin, 103, 516-523.

Harpending R (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology, 66, 591-600.

Hart MW, Sunday J (2007) Things fall apart: biological species form unconnected parsimony networks. Biology Letters, doi:10.1098/rsbl.2007.0307.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences, 270, 313-321.

Heist EJ (2005) Genetics: stock identification. In: Management techniques for elasmobranch fisheries. FAO Fish Tech Pap 474 (eds. JA Musick,R Bonfil), pp. 62-75. United Nations Food and Agriculture Organization, Rome.

Hilborn R, Quinn TP, Schindler DE, Rogers DE (2003) Biocomplexity and fisheries sustainability. Proceedings of the National Academy of Sciences, 100, 6564-6568.

Hudson RR (2000) A New Statistic for Detecting Genetic Differentiation. Genetics, 155, 2011-2014.

ICCAT (2009) Report of the 2009 Porbeagle Stock Assessments Meeting SCRS 2009/014, 1-42.

Jensen C, Natanson L, Pratt HJ, Kohler N, Campana S (2002) The reproductive biology of the porbeagle shark (*Lamna nasus*) in the western North Atlantic Ocean. Fishery Bulletin, 100, 727-738.

Karl SA, Castro ALF, Lopez JA, Charvet P, Burgess GH (2011) Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. Conservation Genetics, 12, 371-382.

Keeney D, Heist E (2006) Worldwide phylogeography of the blacktip shark (*Carcharhinus limbatus*) inferred from mitochondrial DNA reveals isolation of western Atlantic populations coupled with recent Pacific dispersal. Molecular Ecology.

Kohler N, Turner P, Hoey J, Natanson L, Briggs R (2002) Tag and recapture data for three pelagic shark species: blue shark (*Prionace glauca*), shortfin mako (*Isurus*

oxyrinchus), and porbeagle (*Lamna nasus*) in the north Atlantic ocean. Collective Volume of Scientific Papers. International Commission for the Conservation of Atlantic Tunas [Collect. Vol. Sci. Pap. ICCAT]. SCRS/2001/064, 54, 1231-1260.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular genetics analysis and sequence alignment. Briefings in Bioinformatics, 5, 150-163.

Last PR, Stevens JD (2009) Sharks and rays of Australia, 2nd edn. Harvard University Press, Cambridge, Massachusetts.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451-1452.

Maisey JG (1984) Higher elasmobranch phylogeny and biostratigraphy. Zoological Journal of the Linnean Society, 82, 33-54.

Martin AP, Pardini A, Noble L, Jones C (2002) Conservation of a dinucleotide simple sequence repeat locus in sharks. Molecular Phylogenetics and Evolution, 23, 205-213.

Moritz C (1994) Defining evolutionarily significant units for conservation. Trends in Ecology and Evolution, 9, 373-375.

Myers RA, Worm B (2005) Extinction, survival or recovery of large predatory fishes. Philosophical transactions of the Royal Society B, 360, 13-20.

Nance HA, Klimley P, Galvan-Magana F, Martinez-Ortiz J, Marko PB (2011) Demographic Processes Underlying Subtle Patterns of Population Structure in the Scalloped Hammerhead Shark, *Sphyrna lewini*. PLOS One, 6, e21459.

Natanson L, Mello J, Campana S (2002) Validated age and growth of the porbeagle shark, *Lamna nasus*, in the western North Atlantic Ocean. Collective volume of scientific papers. International Commission for the Conservation of Atlantic Tunas, 54, 1261-1279.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a markov chain monte carlo approach (MDIV). Genetics, 158, 885-896.

Pade NG, Queiroz N, Humphries NE, Witt MJ, Jones CS, Noble LR, Sims DW (2009) First results from satellite-linked archival tagging of porbeagle sharks, *Lamna nasus*: Area fidelity, wider-scale movements and plasticity in diel depth changes. Journal of Experimental Marine Biology and Ecology, 370, 64-74.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration and secondary contact. Evolution, 56, 1776-1788.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics, 14, 817-818.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol., 20, 76-86.

Reif WE, Saure C (1987) Shark biogeography: vicariance is not even half the story.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution, 9, 552-569.

Saunders RA, Royer F, Clarke MW (2010) Winter migration and diving behaviour of porbeagle shark, *Lamna nasus*, in the Northeast Atlantic. ICES Journal of Marine Science, 68, 166-174.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. Genetics, 152, 1079-1089.

Schultz JK, Feldheim KA, Gruber SH, Ashley MV, McGovern TM, Bowen BW (2008) Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). Molecular Ecology, 17, 5336-5348.

Shaklee JB, Currens KP (2003) Genetic stock identification and risk assessment. In: Population genetics: principles ans applications for fisheries scientists (ed. EM Hallerman), pp. 291-328. American Fisheries Society, Bethesda, Maryland.

Shivji M, Clarke S, Pank M, Natanson L, Kohler N, Stanhope M (2002) Genetic identification of pelagic shark body parts for conservation and trade monitoring. Conservation Biology, 16, 1036-1047.

Stevens J (1990) Further results from a tagging study of pelagic sharks in the Northeast Atlantic. Journal of the Marine Biological Association of the United Kingdom, 70, 707-720.

Stevens J, Bonfil R, Dulvy N, Walker P (2000) The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. ICES Journal of Marine Science, 57, 476-494.

Stevens J, Fowler SL, Soldo A, McCord M, Baum J, Acuna E, Domingo A, Francis M (2009) *Lamna nasus*. In: 2009 IUCN Red List of Threatened Species. Version 2009.1. http://www.iucnredlist.org/details/11200/0. Downloaded on 04 June 2009.

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics, 143, 1457-1465.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Moleclar Biology and Evolution, 10, 512-526.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. The Journal of Heredity, 89, 438-450.

Wong EH-K, Shivji MS, Hanner RH (2009) Identifying sharks with DNA barcodes: assessing the utility of a nucleotide diagnostic approach. Molecular Ecology Resources, 9, 243-256.

Worm B, Sandow M, Oschlies A, Lotze HK, Myers RA (2005) Global patterns of predator diversity in the open oceans. Science, 309, 1365-1369.

Wright S (1931) Evolution in mendelian populations. Genetics, 16, 97-159.

Wright S (1951) The genetical structure of populations. Annals of Eugenics, 15, 323-354.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, 19, 395-420.

Tables

Table 1. Population statistics for *Lamna nasus*. Number of individuals (*n*), number of haplotypes (*nh*), haplotype diversity (*h*), standard deviation (*SD*), nucleotide diversity (π).

Collection site	*n*	*nh*	h +/- SD	π +/- SD
North Atlantic (NAtl)				
Western Natl (WNA)	40	37	0.996 +/- 0.007	0.0070 +/- 0.0037
Denmark (DK)	30	24	0.979 +/- 0.016	0.0072 +/- 0.0039
United Kingdom (UK)	5	5	1.000 +/- 0.127	0.0089 +/- 0.0058
Total North Atlantic	75	54	0.987 +/- 0.005	0.0071 +/- 0.0037
Southern Hemisphere				
Chile (CHL)	49	19	0.860 +/- 0.034	0.0027 +/- 0.0016
Falkland Islands (FKI)	13	5	0.731 +/- 0.096	0.0036 +/- 0.0022
South Africa (SAfr)	31	15	0.843 +/- 0.050	0.0030 +/- 0.0018
New Zealand (NZ)	51	21	0.853 +/- 0.036	0.0028 +/- 0.0016
Tasmania (TAS)	5	5	1.000 +/- 0.127	0.0032 +/- 0.0023
Total S Hemisphere	149	46	0.843 +/- 0.022	0.0028 +/- 0.0016
Total	224	100	0.929 +/- 0.012	0.0133 +/- 0.0066

Table 2. *L. nasus* AMOVA, SAMOVA, and nearest neighbor statistics: the amount of genetic diversity within and among groups.

Table 2a. Pairwise Φ_{ST} and *Snn* (nearest neighor statistics). Pairwise Φ_{ST} are below the horizontal, pairwise *Snn* are above the horizontal and italicized. Values in bold indicate significance at $P < 0.05$ after sequential Bonferroni correction.

	WNA	ENA	SAfr	NZ	CHL	FKI
Western North Atlantic (WNA)		*0.46*	***1.00***	***1.00***	***1.00***	***1.00***
Eastern North Atlantic (ENA)	-0.02		***1.00***	***1.00***	***1.00***	***1.00***
South Africa (SAfr)	**0.79**	**0.79**		*0.49*	*0.47*	*0.55*
New Zealand (NZ)	**0.81**	**0.81**	-0.02		*0.48*	*0.67*
Chile (CHL)	**0.81**	**0.81**	-0.02	-0.02		*0.68*
Falkland Islands (FKI)	**0.76**	**0.76**	-0.02	-0.01	-0.01	

Table 2b. Heirarchical AMOVA.

		Φ Statistic	
	% variation	Value	*P* value
Among populations $\boldsymbol{\Phi}_{CT}$	82.7%	0.83	0.0440
Among collection sites within populations Φ_{SC}	-0.4%	-0.02	0.9893
Within collection sites Φ_{ST}	17.6%	0.82	0.0000

Table 2c. SAMOVA, K=2.

		Φ Statistic	
	% variation	Value	*P* value
Among groups $\boldsymbol{\Phi}_{CT}$	82.3%	0.82	0.0000
Among collection sites within groups Φ_{SC}	-0.3%	-0.02	0.0062
Within collection sites Φ_{ST}	18.0%	0.82	0.0000

Table 3. Demographic statistics for *L. nasu*s. Fu's F_S test (F_S), Harpending's Raggedness index (*Hri*), sum of squared differences from mismatch analyses (*SSD*), θ at time 0 (θ_0), θ at present time (θ_1), tau (τ), time since expansion in years before present (Time). Values in bold indicate significance at $P < 0.025$ for Fu's F_S test and $P < 0.05$ for global *SSD* after sequential Bonferroni correction.

Collection site	F_S	*Hri*	*SSD*	Θ_0	Θ_1	τ	Time
North Atlantic (NAtl)							
Western Natl (WNA	**-25.008**	0.006	0.004	0.023	26.797	9.059	3.71E+06
Denmark (DK)	**-12.142**	0.010	0.003	0.515	33.240	8.249	3.37E+06
United Kingdom (UI	-0.421	0.100	0.059	0.004	66.318	13.486	5.52E+06
Total North Atlantic	**-24.913**	0.004	0.003	0.297	26.274	8.814	3.61E+06
Southern Hemisphere							
Chile (CHL)	**-9.443**	0.029	0.012	0.000	6.470	4.180	1.71E+06
Falkland Islands (FK	1.626	0.307	0.129	0.000	4.927	5.031	2.06E+06
South Africa (SAfr)	**-6.186**	0.057	0.025	0.000	5.483	4.775	1.95E+06
New Zealand (NZ)	**-11.494**	0.036	0.015	0.002	6.523	4.359	1.78E+06
Tasmania (TAS)	-2.004	0.080	0.020	0.044	55.029	3.924	1.60E+06
Total S Hemisphere	**-26.379**	0.037	0.016	0.000	5.869	4.414	1.81E+06

Table 4. *L. nasus* MDIV, BEAST, and MIGRATE results. MDIV: Reported values are divergence time in years (t), migration rates per generation (m), effective population size (N_e), and time since most recent common ancestor (TMRCA). BEAST: Time since most recent common ancestor (TMRCA). MIGRATE: reported values are migrants per generation, with m2➔1 indicating migrats from the second population into the first, and vice versa; and the effective population size of each population (N_{e1} and N_{e2}). Collection areas are coded as (1) North Atlantic (NAtl): western North Atlantic (WNA), eastern North Atlantic (ENA); and (2) Southern Hemisphere (SH): Chile (CH), New Zealand (NZ), and South Africa (SAfr).

Comparison	MDIV t (years)	MDIV m (per gen)	MDIV N_{ef}	MDIV TMRCA (years)	Beast TMRCA (years)	Migrate m2->1 (per gen)	Migrate m1->2 (per gen)	Migrate N_{ef1}	Migrate N_{ef2}
NAtl - SH	2.73E+06	0.08	2.34E+06	3.84E+06		0.78	0.45	2.26E+06	1.23E+06
WNA - ENA	4.45E+03	168	1.11E+06	1.69E+06		35	55	1.97E+06	1.34E+06
CHL - NZ	3.54E+03	77	4.42E+05	9.71E+05		31	84	6.00E+05	---
CHL - SAfr	4.01E+03	64	3.34E+05	9.83E+05		17	130	---	4.34E+05
NZ - SAfr	1.13E+03	135	4.25E+05	1.30E+06		26	105	7.24E+05	---
North Atlantic (NAtl)					2.87E+06				
Southern hemisphere (SH)					6.39E+05				

Figures

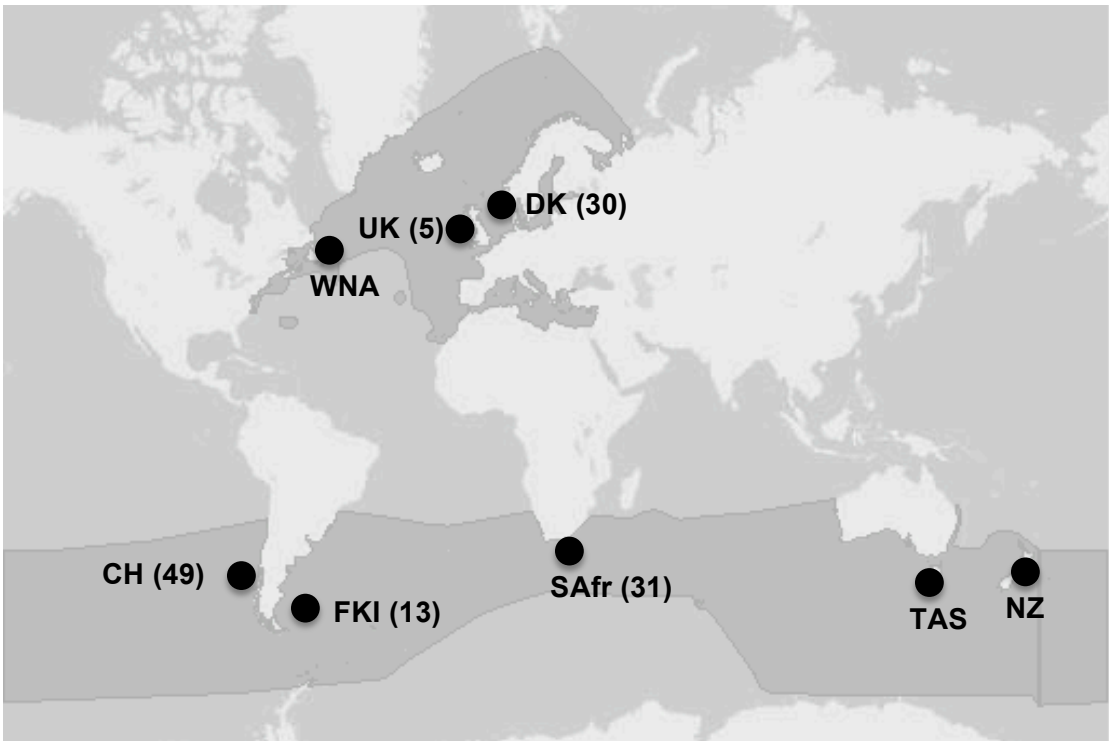


Figure 1: Map of known distribution of *Lamna nasus* (darker gray areas) and sample sizes. Sample numbers for each location are shown in parentheses. Collection areas are coded as follows: western North Atlantic (WNA), United Kingdom (UK), Denmark (DK), Chile (CH), Falkland Islands (FKI), South Africa (SAfr), Tasmania (TAS), and New Zealand (NZ).

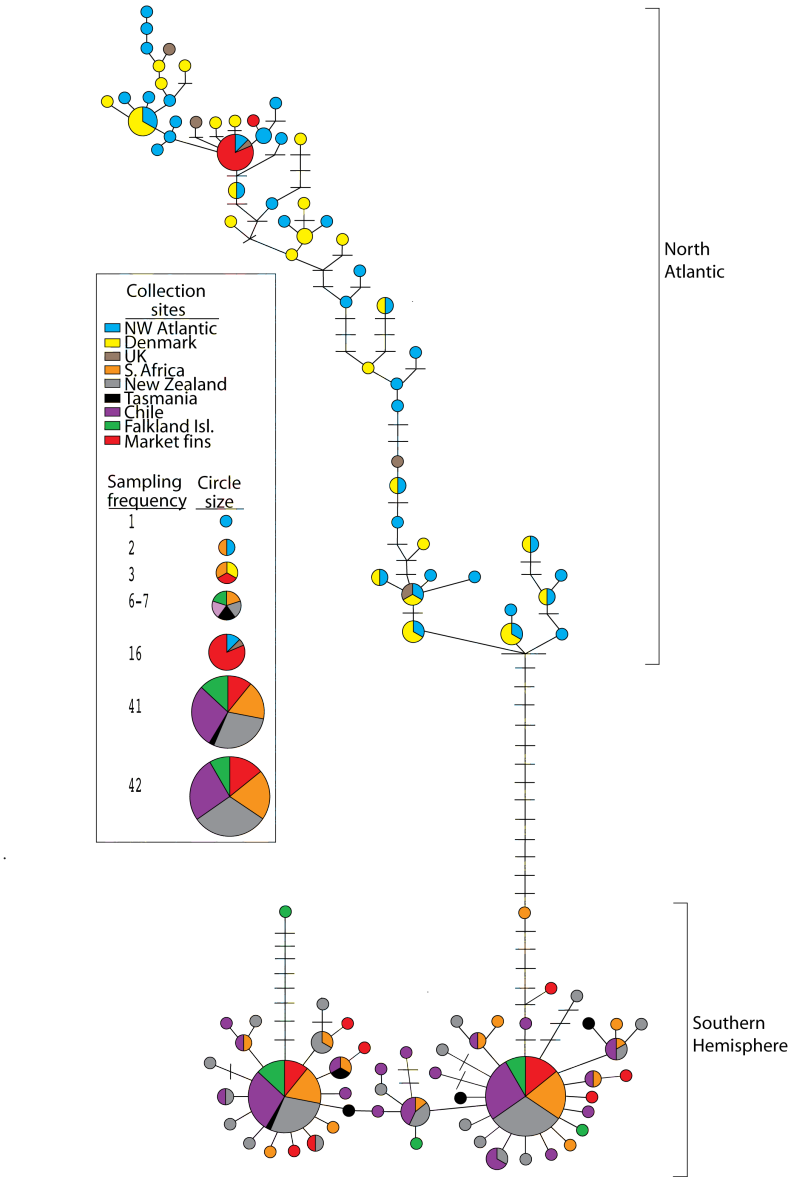


Figure 2. *L. nasus* mtCR statistical parsimony network. Circles represent individual haplotypes with circle size proportional to sampling frequency. Colored pie slices are proportional to the number of individuals from each sampling location with that haplotype. Unbroken connecting lines are equivalent to one mutation step and dashes on connecting lines represent inferred, unsampled haplotypes. The north Atlantic clade at the top of the figure is separated from the southern hemisphere clade at the bottom by 15 mutation steps.

Figure 3a North Atlantic

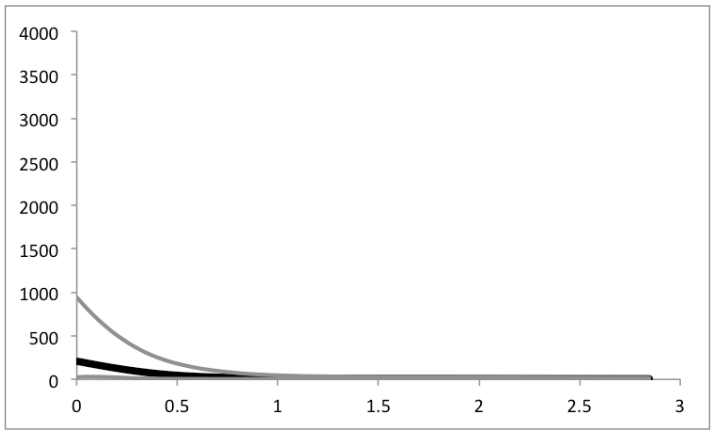
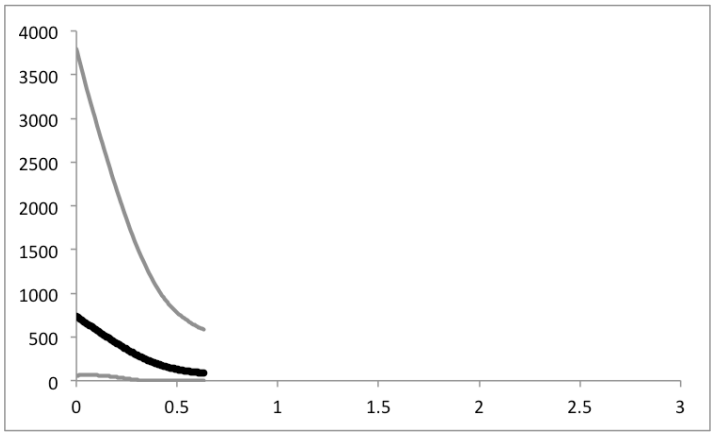


Figure 3b Southern Hemisphere



Female effective population size (thousands) times generation

Time (my)

Figure 3: *L. nasus* Bayesian skyline plots inferred by BEAST based on control region sequences, showing changes over time of the female effective population size times generation time ($N_{ef}T$), of the northern (figure 3a) and southern (figure 3b) populations. Black lines are median estimates of $N_{ef}T$, and gray lines represent the 95% highest posterior density limits.

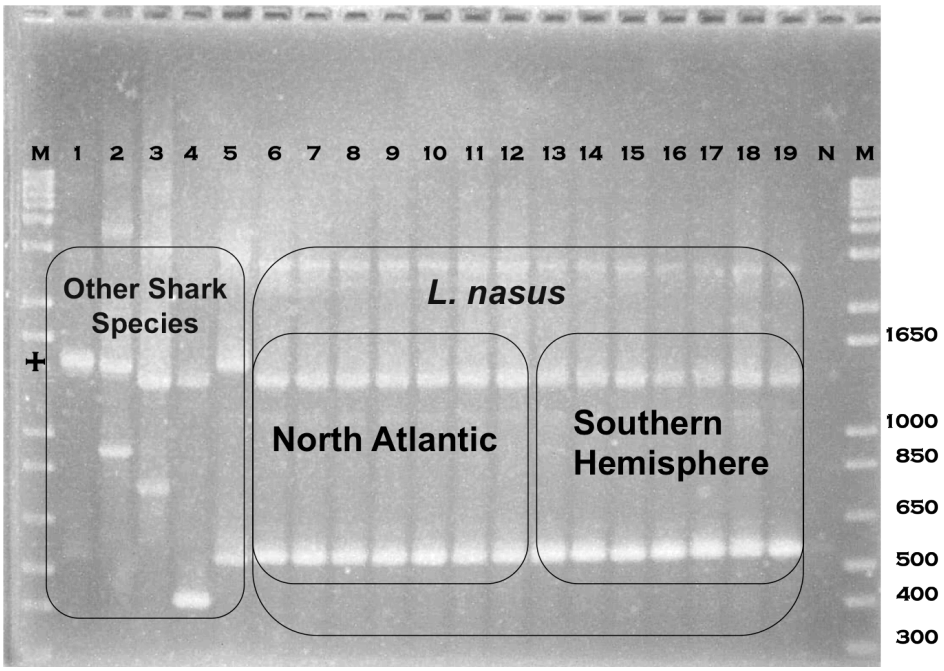


Figure 4: Representative results of PCR amplification with the species ID and hemisphere ID multiplexes. Figure 4a: Testing additional *L. nasus* individuals from known catch locations using species ID multiplex (Shivji et al., 2002). Expected species-specific amplicon sizes given in parenthesis. Lanes 1 - 5 used non-target shark species DNA, lanes 6-12 used North Atlantic *L. nasus* DNA (554 bp) and lanes 13 - 19 used Southern Hemisphere *L. nasus* DNA (554 bp). Non-target species: 1. *C. falciformis* (1 085 bp); 2. *P. glauca* (929 bp); 3. *I. oxyrinchus* (771 bp); 4. *I. paucus* (410 bp); 5. *C. obscurus* (480 bp with ~ 1 476 bp positive control). Expected positive control amplicon is indicated by + in left-hand M lane and was approximately 1 353 bp for Lamniformes and 1 476 bp for Carcharhiniformes. Lanes labeled N contain the negative control reaction (no shark DNA). Lanes labeled M contain the molecular size standard**.** Sizes of the individual size standard bands are indicated to the right of the picture.

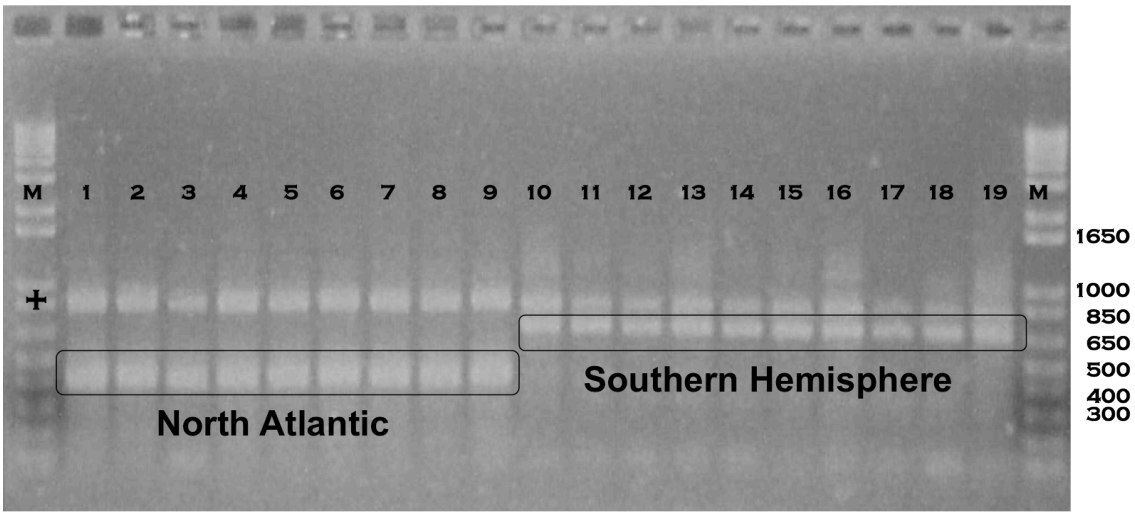


Figure 4b. Testing *L. nasus* individuals from known catch locations using hemisphere ID multiplex. Expected amplicon sizes given in parenthesis. Lanes 1 - 9 used North Atlantic *L. nasus* DNA (400 bp) and lanes 10 - 19 used Southern Hemisphere *L. nasus* DNA (650 bp). Positive control amplicon (800 bp) is indicated by + in left-hand M lane. The negative control reaction (no shark DNA) for this set of PCR was blank as expected and is not shown. All other annotations are as in figure 4a.

Supplementary Information

Table S1. *Lamna nasus* amplification and sequencing primers.

Primer Name	Primer Sequence (5'-3')	Locus	Reference
Control Region amplification and sequencing			
CRF6	AAGCGTCGACCTTGTAAGTC	tRNA-thr	This study
DasR2	GCTGAAACTTGCATGTGTAA	12s rRNA	V. Richards, unpublished
LnasCRF8	TTGACCAGACCTGGCATCTG	Control region	This study
LnasCRR5	CGCGAATGATGAGTACTGAG	Control region	This study
CB3R-LF	CATATTAAACCCGAATGATATTT	CytB	Palumbi 1996
12SA-H1067R	ATAATAGGGTATCTAATCCTAGTTT	12s rRNA	Martin *et al.*, 1992
ITS2 amplification and sequencing, Species ID			
Fish58SF	TTAGCGGTGGATCACTCGGCTCGT	58s rRNA	Shivji *et al.*, 2002
Fish28SR	TCCTCCGCTTAGTAATATGCTTAAATTCAGC	28s rRNA	Shivji *et al.*, 2002
Lnas556F	GGGCACACAGAGGGAGGTTTG	ITS2	This study
Lnas827F	GTCGTCGGCGCCAGCCTTCTA AC	ITS2	Shivji *et al.*, 2002
Lnas485R	TCTCTTTCGGCCGATGCACTC	ITS2	This study
Slmn739R	GCACTCACGAACAGTATGCCA	ITS2	Abercrombie, 2004
Hemisphere-specific primers			
Lnas125F	AATACAAGGGCATATCTCATCTCGACTACATTACAAT	Control region	This study
Lnas865R	TACAAAGCAGGGGGAAGTCTAATAACAGTAAATTTA	Control region	This study
Lnas556F-3	CCTTCGTCCTTGATCGCGTCAAGATTTATTTTCCACCCTGCTTTTTTC	Control region	This study
Lnas293F-2	CAGTCCCCATTAACCTATAATCAAGATCTCCATTTCATAA	Control region	This study

Table S2. *L. nasus* haplotype distribution.

Haplotype	WNAtl	Denmark	UK	South Africa	New Zealand	Tasmania	Chile	Falkland Isl.	# Indivs
1	1								1
2	1	1							2
3	1								1
4	1								1
5	1								1
6	1	1							2
7	1	1							2
8	2								2
9	1								1
10	1								1
11	1								1
12	1								1
13	2		1						3
14	1								1
15	1								1
16	1								1
17	1								1
18	1	1							2
19	1								1
20	2	4							6
21	1								1
22	1								1
23	1								1
24	1								1
25	1								1
26	1								1
27	1								1
28	1	1							2
29	1								1
30	1								1
31	1	1	1						3
32	1								1
33	1								1
34	1	2							3
35	1	1							2
36	1	2							3
37	1								1
38		1							1
39		1							1
40		2							2
41		1							1
42		1							1
43		1							1
44		1							1
45		1							1
46		1							1
47		1							1
48		1							1
49		1							1
50		1							1
51		1							1
52			1						1
53			1						1
54			1						1
55				8	13	1	13	6	41
56				1					1
57				10	15		13	4	42
58				1		1	1		3
59				1			1		2
60				1					1
61				1					1
62				1	2				3
63				1	2		3		6
64				1					1
65				1					1
66				1			1		2
67				1			1		2
68				1	3		3		7
69				1					1
70					1				1
71					1				1
72					1				1
73					1				1
74					1				1
75					1				1
76					1		2		3
77					1				1
78					1				1
79					1				1
80					1		1		2
81					1				1
82					1				1
83					1				1
84					1				1
85					1				1
86						1			1
87						1			1
88						1			1
89							1		1
90							2		2
91							1		1
92							1		1
93							1		1
94							1		1
95							1		1
96							1		1
97							1		1
98								1	1
99								1	1
100								1	1
	40	30	5	31	51	5	49	13	224

Table S3. Net pairwise genetic distances.

Table S3a. Net pairwise genetic distances based on mitochondrial control region.

	n	π (within group diversity)	*Lnas* NAtl	*Lnas* SH	*L. nasus*	*Ldit*	*Ccar*	*Ioxy*	
Lamnids									
L. nasus N. Atl. (Lnas NAtl)	75	0.0073							
L. nasus S. Hemis. (Lnas SH)	149	0.0042	0.0200						
L. nasus (Lnas)	224	0.0155							
L. ditropis (Ldit)	3	0.0078			0.0512				
C. carcharias (Ccar)	3	0.0266			0.1623	0.1685			
I. oxyrinchus (Ioxy)	4	0.0068			0.1078	0.1223	0.1605		
I. paucus (Ipau)	4	0.0037			0.1095	0.1230	0.1515	0.0743	
Carcharhinids			*Cacr*	*Calb*	*Camb*	*Cbre*	*Cfal*	*Cleu*	*Cper*
C. acronotus (Cacr)	1	N/A							
C. albimarginatus (Calb)	2	0.0000	0.0550						
C. amblyrhynchos (Camb)	3	0.0052	0.0502	0.0359					
C. brevipinna (Cbre)	8	0.0052	0.0437	0.0427	0.0461				
C. falciformis (Cfal)	9	0.0050	0.0390	0.0407	0.0410	0.0420			
C. leucas (Cleu)	8	0.0036	0.0553	0.0402	0.0527	0.0553	0.0535		
C. perezi (Cper)	1	N/A	0.0569	0.0549	0.0575	0.0521	0.0531	0.0515	
C. signatus (Csig)	7	0.0037	0.0507	0.0558	0.0551	0.0551	0.0520	0.0642	0.0569
Sphyrnids			*Eblo*	*Slew*	*Smok*				
E. blochii (Eblo)	1	N/A							
S. lewini (Slew)	2	0.0055	0.1654						
S. mokarran (Smok)	6	0.0077	0.1712	0.1967					
S. zygaena (Szyg)	4	0.0024	0.1382	0.1421	0.1449				

Table S3b. Net pairwise genetic distances based on mitochondrial COI.

Lamnids	n	π (within group diversity)	*Lnas* NAtl	*Lnas* SH	*L. nasus*	*Ldit*	*Ccar*	*Ioxy*	
L. nasus N. Atl. (Lnas NAtl)	30	0.0030							
L. nasus S. Hemis. (Lnas SH)	50	0.0055	0.0069						
L. nasus (Lnas)	80	0.0078							
L. ditropis (Ldit)	3	0.0025			0.0513				
C. carcharias (Ccar)	3	0.0049			0.1539	0.1625			
I. oxyrinchus (Ioxy)	4	0.0073			0.1356	0.1375	0.1522		
I. paucus (Ipau)	4	0.0016			0.1241	0.1372	0.1699	0.1428	
Carcharhinids			*Cacr*	*Calb*	*Camb*	*Cbre*	*Cfal*	*Cleu*	*Cper*
C. acronotus (Cacr)	1	N/A							
C. albimarginatus (Calb)	2	0.0031	0.0433						
C. amblyrhynchos (Camb)	3	0.0021	0.0456	0.0340					
C. brevipinna (Cbre)	8	0.0008	0.0333	0.0428	0.0418				
C. falciformis (Cfal)	9	0.0019	0.0401	0.0353	0.0287	0.0430			
C. leucas (Cleu)	8	0.0018	0.0440	0.0525	0.0514	0.0490	0.0515		
C. perezi (Cper)	1	N/A	0.0382	0.0524	0.0522	0.0398	0.0497	0.0540	
C. signatus (Csig)	7	0.0012	0.0438	0.0506	0.0461	0.0518	0.0451	0.0513	0.0606
Sphyrnids			*Eblo*	*Slew*	*Smok*				
E. blochii (Eblo)	1	N/A							
S. lewini (Slew)	2	0.0000	0.0896						
S. mokarran (Smok)	6	0.0005	0.0768	0.0965					
S. zygaena (Szyg)	4	0.0008	0.0845	0.0936	0.0925				

Table S3c. Net pairwise genetic distances based on nuclear ITS2.

Lamnids			*Lnas* Natl	*Lnas* SH	*L. nasus*	*Ldit*	*Ccar*	*Ioxy*	
L. nasus N. Atl. (Lnas NAtl)	10	0.0000							
L. nasus S. Hemis. (Lnas SH)	9	0.0000	0.0089						
L. nasus (Lnas)	19	0.0049							
L. ditropis (Ldit)	3	0.0083			0.0144				
C. carcharias (Ccar)	3	N/A			0.0608	0.0626			
I. oxyrinchus (Ioxy)	4	0.0000			0.0534	0.0565	0.0682		
I. paucus (Ipau)	4	0.0000			0.0686	0.0749	0.0825	0.0783	

Carcharhinids			*Cacr*	*Calb*	*Camb*	*Cbre*	*Cfal*	*Cleu*	*Cper*
C. acronotus (Cacr)	1	N/A							
C. albimarginatus (Calb)	2	0.0011	0.0500						
C. amblyrhynchos (Camb)	3	0.0003	0.0482	0.0068					
C. brevipinna (Cbre)	8	N/A	0.0400	0.0481	0.0481				
C. falciformis (Cfal)	9	0.0039	0.0535	0.0416	0.0380	0.0491			
C. leucas (Cleu)	8	0.0029	0.0153	0.0440	0.0420	0.0323	0.0453		
C. perezi (Cper)	1	0.0006	0.0454	0.0189	0.0171	0.0445	0.0371	0.0386	
C. signatus (Csig)	7	0.0000	0.0491	0.0518	0.0500	0.0473	0.0501	0.0431	0.0472

Sphyrnids			*Slew*	*Smok*
S. lewini (Slew)	2	0.0006		
S. mokarran (Smok)	6	0.0000	0.0515	
S. zygaena (Szyg)	4	0.0000	0.0344	0.0344

Table S4. Divergence times and rates. The complete mitochondrial control region sequence was used unless otherwise noted.

Family	Species compared	Divergence Time (mya)	Reference	d corr	Divergence (%/my)
Lamnidae	*C. carcharias* vs. *L. ditropis* and *L. nasus*	60-65	Martin et al. (1992)	0.152	0.25
Lamnidae	*C. carcharias* vs. *L. ditropis* and *L. nasus*	50-55	Capetta et al. (1987)	0.152	0.30
Lamnidae	*C. carcharias* vs. *I. oxyrinchus* and *I. paucus*	60	Martin et al. (1992)	0.135	0.22
Lamnidae	*C. carcharias* vs. *I. oxyrinchus* and *I. paucus*	50-55	Capetta et al. (1987)	0.135	0.27
Lamnidae	*I. oxyrinchus* vs. *I. paucus*	25-30	Maisey et al. (1984)	0.074	0.30
Lamnidae	*I. oxyrinchus* and *I. paucus* vs. *L. ditropis* and *L. nasus*	60-65	Martin et al. (1992)	0.081	0.13
Lamnidae	*I. oxyrinchus* and *I. paucus* vs. *L. ditropis* and *L. nasus*	50-55	Capetta et al. (1987)	0.081	0.16
Lamnidae	**Average of Lamnid mutation rates**				**0.23**
Carcharhinidae vs. Sphyrnidae	*C. brevipinna, C. falciformis, C. leucas* and *C. signatus* vs. *S. mokarran* and *S. zygaena*	38	Capetta et al. (1987); Maisey et al. (1984)	0.243	0.64
Carcharhinidae	West Atlantic vs. East Pacific populations of *C. leucas*	3.2		0.012	0.38
Other reported mutation rates					
Sphyrnidae	*S. lewinii,* 548 bp control region		Duncan et al. (2006)		0.80
Sphyrnidae	8 Sphyrnid species, 548 bp control region		Nance et al. (2011)		1.21
Carcharhinidae	*C. limbatus*		Keeney et al. (2006)		0.43
Carcharhinidae	*N. brevirostris* and *N. acutidens*		Schultz et al. (2008)		0.67
Sphyrnidae	*S. tiburo,* mt CytB		Martin & Palumbi (1993)		2.30
	Centropomidae spp. (Snook)		Donaldson & Wilson (1999)		3.60

CHAPTER 2: Assessment of global population genetic structure and genetic diversity in the bull shark *Carcharhinus leucas*

Abstract

Population structure, genetic diversity, and demographic trends of exploited species are important information components of national and international management and conservation efforts. The bull shark (*Carcharhinus leucas*) is a globally distributed, large coastal shark that occurs in marine, estuarine and freshwater habitats. It is caught in recreational and commercial fisheries throughout its range, and it has been assessed as near threatened by the IUCN. Bull sharks show strong mitochondrial population structuring in the two regions examined thus far: between the northern and southern hemisphere in the western Atlantic, and along the northern coast of Australia, but no structure within either region when assessed at a modest number (up to five) of nuclear microsatellite loci. To obtain a global and higher resolution perspective on the population genetic structure and demographics of the bull shark, I assessed genetic variation at eleven microsatellite loci in 468 samples from across the species distribution. The microsatellite data revealed significant genetic partitioning between bull sharks from the western North Atlantic, Indo-Australia and Fiji. Demographic analyses indicate a slight population decline in the western North Atlantic and increases in Fiji and Indo-Australia, with evidence of historic and contemporary migration between the western North Atlantic and Indo-Australia and between Fiji and Indo-Australia, but not between the western North Atlantic and Fiji. These findings have important management implications because they highlight the potential for male mediated gene flow across large distances of continuous coastline, and much lower levels of gene flow across open waters.

Introduction

There is ample evidence that intensive commercial fishing practices and overfishing have had detrimental effects at the ecosystem and species levels (Watson *et al.* 2013). Species with biological characteristics such as late maturity, slow growth, long gestation periods and low fecundity are particularly vulnerable to fishery pressures. These life history

characteristics are common to many shark species and imply that they may be affected by fishing mortality at greater rates than other marine species (Stevens *et al.* 2000; Cortes 2002; Garcia *et al.* 2008; Au *et al.* 2009; Dulvy & Forrest 2010).

Several studies have indicated that large coastal sharks have been overfished (Baum *et al.* 2003; Fowler *et al.* 2005; Shepherd & Myers 2005; Heithaus *et al.* 2007; Myers *et al.* 2007; Baum & Blanchard 2010; Hisano *et al.* 2011). The bull shark, *Carcharhinus leucas*, is a globally distributed coastal species caught in artisanal, commercial and recreational fisheries throughout its range. Depending on the region, it is a target or bycatch species and exploited opportunistically for its meat and fins. The bull shark is a significant component of fisheries in US waters, comprising 11% of the 1978 recreational shark catch in the Gulf of Mexico (Casey & Hoey 1985) and approximately 16% of the catch of large coastal sharks along the US Atlantic seaboard (Branstetter & Burgess 1997).

Bull sharks are found in continental and near-shore waters as well as estuaries and fresh water rivers. Coastal shark species, including the bull shark are also vulnerable to other anthropogenic effects such as habitat degradation, development, and pollution, although the individual impacts of such effects may be difficult to discern. Additionally, coastal species may be limited by their habitat, exhibiting fidelity to discrete locations for feeding, mating, parturition, maturation and migratory routes (Castro 1983; Simpfendorfer & Milward 1993; Heupel *et al.* 2007; Speed *et al.* 2010; Brunnschweiler & Baensch 2011). Tracking studies have indicated significant site fidelity in many regions in both juvenile and adult bull sharks, although some individuals have also exhibited long-range movements (Yeiser *et al.* 2008; Brunnschweiler *et al.* 2010; Carlson *et al.* 2010). The species exhibits predictable movement patterns, with documented seasonal migrations along the US eastern seaboard (Castro 1983), female migrations into estuaries and rivers for parturition, juvenile migrations between estuarine or riverine nursery areas and offshore waters depending on temperature (Snelson *et al.* 1984; Curtis *et al.* 2011; Matich & Heithaus 2012), and an ontogenetic shift in habitat to primarily near-shore waters as the animals mature (Carlson *et al.* 2010). Population demographic studies are notably lacking, although fine scale studies have indicated localized

population increases (Froeschke *et al.* 2012) and decreases (O'Connell *et al.* 2007) in different regions of the Gulf of Mexico.

These complex behavioral patterns of high potential vagility, observed site fidelity, long distance individual movements, and seasonal migrations along with conservation and management information needs have prompted inquiry into population differentiation and genetic diversity in two regional studies of bull sharks to date. Karl *et al.* (2011) found strong matrilineal genetic differentiation between populations in the north- and southwestern Atlantic sequences, but no significant differentiation between these regions with five nuclear microsatellite loci. Their results are consistent with female philopatry and male mediated gene flow, as has been observed in many other sharks (Pardini *et al.* 2001; Feldheim *et al.* 2002; Keeney *et al.* 2005; Schultz *et al.* 2008; Daly-Engel *et al.* 2012; Testerman *et al.* in prep-b). Tillett *et al.* (2012) examined population structure in juvenile bull sharks sampled from 13 river systems across Northern Australia, and also found high levels of matrilineal differentiation but no nuclear differentiation (based on three microsatellite loci), supporting female reproductive philopatry in this species.

I expand on these previous regional studies by analyzing a globally distributed sample set (n = 468) using 11 nuclear microsatellite loci to obtain a higher resolution perspective on bull shark population dynamics. I investigated the extent of nuclear population structure and genetic diversity present in the species, and the population demographics of genetically distinct populations, including long-term effective population size and the amount of contemporary and historical migration between populations. I also evaluated the observed regional and global levels of genetic diversity for signatures of decreased population size that might have resulted from overexploitation or other anthropogenic effects. I highlight the implications of these results for conservation and management efforts of this widespread but still enigmatic species.

Materials and methods

Samples and DNA extraction

Bull shark samples were obtained from other researchers, fisheries observer programs and recreational and artisanal fishers from 16 collection areas distributed in three ocean

basins (Figure 1). Herein, the term Indo-Australia includes the Indian Ocean and Australian collection areas. Tissue samples were obtained as fin clips or muscle punches and preserved in 95% ethanol. Genomic DNA was isolated from tissue samples using the QIAGEN DNeasy Tissue kit (QIAGEN, Inc., Valencia, CA).

Microsatellite genotyping

I assessed the nuclear DNA variability by initially genotyping 605 individuals at 11 previously published nuclear microsatellite loci from other carcharhinid species, with amplification conditions experimentally optimized for bull shark (Supplementary Information, Table S1). Microsatellite loci were amplified using forward primers with fluorescently labeled M13(-21) attached to their 5′ ends (Schuelke 2000). PCR reactions were performed in 12 μl volumes containing 1μl of unquantified genomic DNA, 300 μM of each dNTP, 2 pM forward primer, 5 pM reverse primer, 5 pM M13 tail, 1.5mM $MgCl_2$, 0.5 U HotStar *Taq*™ DNA polymerase (QIAGEN, Inc.) and 1.25μl 10x HotStar *Taq*™ reaction buffer. The few exceptions were 1.8 mM $MgCl_2$ and 200 μM of each dNTP for locus LS24, and 200 μM of each dNTP for locus Clim108. PCR cycling conditions were 95ºC for 15 minutes; followed by 35 cycles of 94ºC for 1 minute, the locus specific annealing temperature (Table S1) for 1 minute, and 72ºC for 1 minute, with a final extension of 5 minutes at 72ºC. Fragments were separated on an AB 3130 genetic analyzer (Applied Biosystems, Inc., Foster City, CA). Genotypes were scored using GENEMAPPER software version 3.7 (Applied Biosystems, Inc.) by comparison with the internal size standard LIZ 500. Approximately 5% of the genotypes were reamplified and rescored to ensure genotyping repeatability and quality.

Individual samples that failed to amplify at more than 3 loci were excluded from further analyses. Since I received samples from some of the same collection areas at different times, sample genotypes were checked for duplicates using MICROSATELLITE TOOLKIT (PARK 2001) for EXCEL. Samples with identical multilocus genotypes or with up to two allele mismatches were considered to potentially be from the same individual; these duplicate sample genotypes were re-evaluated and corrected as appropriate. Several juveniles and small individuals of unknown age were sampled at various collection areas. Because juvenile bull sharks may use the same nursery site for a few

years (Wiley & Simpfendorfer 2007; Heithaus *et al.* 2009), it is possible that these juveniles may have been related; I used the pairwise hypothesis testing option in ML-RELATE (Kalinowski *et al.* 2006) to identify related individuals (full sibs, half sibs, or parent-offspring) in each collection area. Likelihood ratios of 10 000 random dyads were simulated, and the hypothesis that a pair was related was accepted when the probability of their likelihood ratio was < 0.01. When duplicate or related individuals were found, one individual of each pair was randomly selected for inclusion in further analyses. All sample genotypes were checked for null alleles, scoring errors, and large allele dropout using MICROCHECKER (Van Oosterhout *et al.* 2004). MICROCHECKER was also used to generate adjusted genotypes corrected for the presence of null alleles. These adjusted genotypes were used to obtain corrected F_{IS} and F_{ST} values.

Analysis of genetic diversity and population structure

Input files were prepared for microsatellite data analyses using the software CREATE v. 1.36 (Coombs *et al.* 2008). Microsatellite loci were checked for deviations from Hardy-Weinberg equilibrium and for evidence of linkage disequilibrium using GENEPOP v4.1 (Rousset 2008). Exact tests of Hardy-Weinberg equilibrium were performed using the Markov chain method (Guo & Thompson 1992; Rousset & Raymond 1995), with 10 000 dememorizations, 20 batches and 5 000 iterations per batch.

The software GENETIX v4.05.2 (Belkhir *et al.* 1996) was used to estimate observed (H_o) and expected (H_e) levels of heterozygosity, number of alleles, and inbreeding coefficients (F_{IS}). I used FSTAT v2.9.3.2 (Goudet 1995, 2001) to calculate genetic diversity (Nei 1987) and allelic richness (El Mousadik & Petit 1996). The frequency of null alleles was estimated using the Expectation Maximization (EM) algorithm of Dempster *et al.* (1977) as implemented in GENEPOP.

Due to small sample sizes from the Philippines (n=1), New Caledonia (n=1) and the eastern Pacific (n=5), genotype data from these collection areas were not used in population-level statistical analyses, but were included in individual-based analyses. For all statistical analyses with multiple tests, I evaluated significance levels ($P < 0.05$) after sequential Bonferroni correction. I used GENEPOP to test for genic differentiation by calculating pairwise F_{ST} values among the 13 collection areas with $n \geq 10$ samples. Fisher's exact tests for significance of genic differentiation were performed using the

Markov chain method, with 10 000 dememorizations, 20 batches and 5 000 iterations per batch. Global tests across loci and collection areas used Fisher's method to calculate overall significance values. Additionally, an unbiased estimator of divergence, Jost's D_{est} was calculated using SMOGD (CRAWFORD 2010). Neighboring collection areas that were not significantly differentiated from each other based on F_{ST} values were pooled into putative populations for subsequent demographic and individual assignment analyses.

I assessed individual-based genetic population structure using three different approaches. First, I assigned individuals to clusters using the Bayesian clustering algorithm in STRUCTURE version 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003). STRUCTURE analyses were performed using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation. Hubisz *et al.* (2009) showed the LOCPRIOR model improved clustering performance in situations where genuine but weak population structure existed. I ran STRUCTURE with the collection areas as prior information and a burn-in period of 100 000 MCMC generations followed by 200 000 iterations. I varied the number of populations from $K = 1$ through $K = 10$ with 10 replicates for each K, using the admixture model and correlated allele frequencies. STRUCTURE HARVESTER v0.6.93 (Earl & vonHoldt 2012) was used to determine the most likely number of distinct genetic clusters by evaluating the logarithm of the probability of the data (lnP(D | K)) (Pritchard *et al.* 2000) and estimates of ΔK (Evanno *et al.* 2005). Each individual's admixture proportions were averaged over the 10 replicates for the most likely K using the program CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007), and the output graphically displayed by the program DISTRUCT v1.1 (Rosenberg 2004). Second, I used a non-MCMC algorithm, FLOCK v2.0 (Duchesne & Turgeon 2009; Duchesne & Turgeon 2012), which uses an iterative reallocation method to assign individuals into K distinct populations without any *a priori* sample location information. FLOCK v2.0 was run starting with $K = 2$ and increasing until one stopping condition was reached, with 50 runs per K and 20 iterations per run. Both programs were used because STRUCTURE may have more power to detect population structure when migration rates are low while FLOCK may have more power under conditions of high, sustained migration (Duchesne & Turgeon 2012). Finally I performed principal coordinate analyses (PCoA) in GENEALEX to visualize any genetic

partitioning present among individuals. Eigenvectors were calculated from a covariance matrix and the first two coordinates were plotted.

Population demographics

Population demographics for the genetically differentiated populations identified by pairwise F_{ST} analyses were assessed using both statistical and coalescent-based methods. Given high exploitation rates and indications of population declines in some regions (Heithaus *et al.* 2007; O'Connell *et al.* 2007), I tested for evidence of bottlenecks in each of the genetically distinct populations by using two statistical tests, the sign-rank test and the M-ratio test. The program BOTTLENECK v1.2.02 (Piry *et al.* 1999) calculates statistics that test for departures from equilibrium patterns of heterozygosity that can be disrupted when the effective population size is significantly reduced. I used the Wilcoxon sign-rank test because it is the most powerful when fewer than 20 loci are analyzed (Piry *et al.* 1999). Calculations were performed assuming the infinite alleles model (IAM), the single-step model (SSM), and the two phase model (TPM) with variance set to 12 and 95% single step mutations as recommended by the authors. However, the IAM is most appropriate for allozymes (Piry *et al.* 1999), therefore I place more weight on the results from the SMM and TPM models. Significance was assessed over 10 000 replicates.

It is expected that during a population bottleneck the number of alleles will decline faster than the range of allele sizes. I conducted a second statistical analysis, the *M*-ratio test, using the software *M_P_val* (Garza & Williamson 2001) and compared the empirical M-ratio to critical values (M_C) obtained from simulated equilibrium distributions calculated in the program *critical_M,* using the two-phase model of microsatellite mutation, 10 000 replicates, with the percentage of mutations that follow the single step model ($p_s = 0.9$) and the mean size of larger mutations ($\Delta_g = 3.5$) as suggested by the authors. Critical values (M_C) were calculated for a range of θ (0.01, 0.1, 1 and 10), and statistical significance was assessed over 10 000 replicates.

I then evaluated changes in effective population size N_e using two coalescent-based methods. First I used the hierarchical Bayesian MCMC model implemented in MSVAR 1.3 (Storz & Beaumont 2002) to estimate parameters including the current population size (N_0), the ancestral population size (N_1), the time of population size

change (t), and the mutation rate (μ). I applied the exponential growth model, five replicate runs with varied priors and broad hyperpriors for the model parameters, and run lengths of 9×10^8 steps with output reported every 20 000 steps. After removing a 10% burn-in, convergence was assessed by calculating the Gelman-Rubin multivariate scale reduction factor (PRSF) and the effective sample size (ESS) determined across the five independent runs using the CODA package in the software R v2.15.1 (R Development Core Team 2012). Estimates of the mode and 95% highest posterior density (HPD) intervals were obtained using the modeest v1.14 and coda v0.15-2 packages, respectively, in R.

The Bayesian implementation in LAMARC v2.1.8 (Kuhner 2006), a coalescent-based MCMC method, was also used to estimate Θ ($\Theta = 4N_e\mu$) and growth rates for genetically differentiated populations as well as directional migration rates ($m = \Theta * M$) between population pairs. Initial data runs were conducted with wide priors to determine appropriate upper bounds and the priors were revised for final data runs. Each data set was analyzed at least three times with different random seeds and run length of 2.0×10^7 generations, while the final input prior values and all other inputs were held constant to evaluate convergence. For each of the output parameters, the means from each of the final runs were averaged to obtain the best estimate.

Individual assignment

I analyzed the number of first generation migrants using the Bayesian individual assignment method as implemented in GENECLASS 2.0 (Piry *et al.* 2004) at $p < 0.05$. I used L=L_Home as not all possible source populations were sampled, the Rannala and Mountain (1997) Bayesian criteria and Monte-Carlo resampling (Paetkau *et al.* 2004) with 10 000 simulated individuals ($\alpha = 0.05$) to assess probabilities. Additionally, I assessed the number of first, second and third generation migrants using the prior population information and Gensback = 2 options in STRUCTURE with the run parameters described above except that I fixed K = 3 based on the Atlantic, Indo-Australia, and Fiji populations defined by F_{ST} and Jost's D analyses. When the posterior probability of an individual shark being from its assumed population was < 0.8, it was assigned to the class with the highest posterior probability, either as a migrant itself or having at least one migrant parent (F1) or grandparent (F2).

Results

Genetic variation and geographic structure

After deleting 23 samples that amplified at less than eight loci, 20 duplicate genotypes and 94 related samples, I obtained the multilocus genotypes for a total of 468 bull shark samples at 11 microsatellite loci. There was no evidence of scoring errors or large allele dropout. Summary statistics including the number of alleles, allelic richness, genetic diversity, expected and observed heterozygosity, inbreeding coefficient, and null allele frequencies are reported in Table 1. Polymorphism ranged from 8 to 41 alleles per locus. Loci CS02, CS07 and CT05 exhibited significant $F_{IS,}$ probably due to the presence of null alleles. The recalculated F_{IS} and F_{ST} values based on genotypes corrected for the presence of null alleles in MICROCHECKER were not significantly different from the uncorrected F_{ST} values based on Mann-Whitney U tests, so uncorrected values are reported for all loci. A few other loci deviated from HWE in one (LS24 and CL102) or two (CL107) collection areas, none deviated significantly from HWE within a genetically differentiated population. A few collection areas showed evidence of linkage disequilibrium at some pairs of loci, but none of the loci showed evidence of linkage disequilibrium consistently across collection areas. Although diversity indices varied across loci, most loci had slightly higher levels of allelic richness and expected heterozygosity in the Indo-Pacific than in the Atlantic.

The global F_{ST} is large (0.14) and significant ($P < 0.001$). Pairwise F_{ST} values between collection areas within the Atlantic, within Indo-Australia, and within Fiji were small and non-significant (values not shown), while pairwise F_{ST} values were large and significant for comparisons between these regions (Table 2). Similarly, Jost's D_{est} values were small for comparisons within these regions but large in comparisons between regions (Table 2). Given the absence of significant population structure within each of the western Atlantic, Indo-Australia and Fiji regions, samples within each region were pooled and considered to comprise three genetically distinct populations for further analyses.

Bayesian STRUCTURE analyses (Figure 2) revealed a maximum log likelihood of posterior probability at $K = 3$ genetically distinct clusters, and the ΔK value also showed a distinct maximum at $K = 3$. However, the FLOCK analyses indicated $K = 2$ genetically

distinct clusters comprising the western Atlantic and the Indo-Pacific regions. Separate hierarchical STRUCTURE and FLOCK analyses of samples from only within the Atlantic, Indo-Australia and Fiji all returned $K = 1$. PCoA analysis (Figure 3) indicates the presence of two overlapping clusters, one in the Atlantic and one comprising the Indo-west Pacific individuals. Surprisingly, the five bull shark individuals from the eastern Pacific clustered with the Atlantic individuals in both STRUCTURE and PCoA analyses.

Demographic analyses

Results of the bull shark demographic analyses are presented in Table 3 and Figure 4. The Wilcoxon test for heterozygosity excess in BOTTLENECK (Table 3a) was not significant ($P > 0.05$) for any of the three populations under any of the models, with the exception of the Fiji population ($P < 0.05$) under the IAM. In contrast, the Wilcoxon test for heterozygosity deficit was significant for the Atlantic and IndoAustralian populations under the TPM indicating population expansions, but non-significant in the Fijian population. The M-ratio analyses (Table 3a) resulted in observed M values smaller than the simulated critical values when $\Theta \leq 1$ for the Atlantic population and $\Theta \leq 0.1$ for the Indo-Australian and Fijian populations.

The MSVAR analyses (Table 3a) indicated a potential decline in the western Atlantic and population increases in the Indo-Australian and Fijian populations, although the 95% highest posterior densities for population sizes and time since population expansion or decline were quite broad. The average estimated mutation rate across all three populations was 3.57×10^{-6}. MCMC diagnostics indicated that the chains were well mixed (PRSF < 1.3) and effective sizes were large (ESS > 200 for each parameter in each population).

Growth rates estimated in LAMARC (Table 3b) were small or negative, and the confidence intervals were close to or spanned 0 indicating a potential population decline (Kuhner 2006). Point estimates of population size derived from LAMARC analyses were smaller than those from the MSVAR analyses (Figure 4). Migration rates between the western Atlantic and both the Indo-Australian and Fijian populations were low with no significant directional bias. However, the migration rate from Indo-Australia into Fiji was twice as large as all the other rates.

GENECLASS identified five first generation migrants: four individuals sampled in Fiji as likely first generation migrants from Indo-Australia, and one individual sampled in South Africa that was likely a migrant from the Atlantic (Table 4). STRUCTURE also identified that same individual sampled in South Africa as a migrant from the Atlantic. The STRUCTURE analyses indicated that one individual sampled in South Africa was a likely migrant from Fiji (posterior probability 0.89). STRUCTURE also identified two individuals that were sampled in South Africa as likely first (approximately 0.6 posterior probability) or second (posterior probability 0.06 - 0.11) generation migrants from Fiji.

Discussion

Patterns of genetic diversity and population structure

Most of the samples used in this study were obtained opportunistically from fisheries. However, the samples from the Everglades were from a study of neonate and juvenile bull sharks in a nursery area (Heithaus *et al.* 2009; Matich & Heithaus 2012; Matich & Heithaus 2014), thus it is not surprising that I found 37 pairs of related individuals out of an initial 138 young sharks genotyped. In other collection areas information on size or age class was generally not available. Additionally, there were 17 pairs of related individuals in a total of 104 bull sharks sampled in a multi-year study in Fiji. The highest percentages of related individuals were in the Everglades (0.27), Florida Keys (0.20), and Fiji (0.16), followed by the western North Atlantic (0.12) and eastern (0.14) and western (0.15) Australia. The proportions of related individuals were lowest in the remaining Indian Ocean sites (0.00 – 0.05). All duplicate genotypes obtained were from pairs of sharks sampled in the same general location and were assumed to represent individuals that were sampled more than once. What is striking is that in Fijian sharks sampled over four years, the genotypes revealed that nine sharks were sampled twice, one shark was sampled three times, and one shark was sampled five times. This resampling at different times supports the site fidelity of bull sharks at the Fiji collection area reported by Brunnschweiler *et al.* (2010) and Brunnschweiler and Baensch (2011). However, as noted by those authors, it is not clear whether this site fidelity is due to the relative isolation of the Fiji island archipelago from nearby suitable coastal habitats in Indo-Australia or to eco-tourism shark feeding activities that may result in artificial site

fidelity. However, the genetic differentiation between Fiji and Indo-Australia is significant after related individuals and duplicate genotypes were removed from the dataset, so it is unlikely that the genetic differentiation is due only to the feeding activity. Additional genetic and telemetry studies are needed to fully resolve this issue.

The levels of microsatellite-based genetic diversity are high compared to the diversity levels reported in other elasmobranchs (summarized in Karl *et al.* (2011). The expected heterozygosity value reported here (H_e = 0.83) is very similar to that reported for western Atlantic bull sharks (H_e = 0.84) by Karl *et al.* (2011), and somewhat larger than that reported for juvenile bull sharks in northern Australia (H_e = 0.80) by Tillett *et al.* (2012). Additionally, it is not surprising that I report a higher average number of alleles (N_a=27) than either of these previous bull shark studies (N_a=19 in Karl *et al.,* 2011; and N_a=13 in Tillet *et al.*, 2012) since this study incorporated more than twice as many loci as well as samples from multiple, globally distributed, genetically distinct populations. However, it is quite unexpected that the average number of alleles reported here is larger than any of the comparable values summarized by Karl *et al.* (2011). Overall, large effective population sizes for all three genetically distinct populations of bull sharks may have contributed to the high levels of genetic diversity.

Many authors have reported distinct nuclear population genetic structure between Atlantic and Indo-Pacific populations of elasmobranchs with little intra-basin structure (Benavides *et al.* 2011; Taguchi *et al.* 2011; Cunha *et al.* 2012; Daly-Engel *et al.* 2012; Testerman *et al.* In prep-a; Testerman *et al.* in prep-b). Similarly, there is significant population structure between the Atlantic and Indo West Pacific sites. Surprisingly, there is also significant structure between samples from Fiji and Indo-Australia as well as between Fiji and the Atlantic. However, the collection areas in the Indo-Australian region are connected by relatively continuous coastline although the distances between them are quite large. This is similar to the findings of no nuclear population structure between the western North and South Atlantic Karl *et al.* (2011) which covers a comparable expanse of near shore waters. In contrast, samples from Fiji and eastern Australia are genetically differentiated although the linear geographic separation is much smaller. Migrants between Fiji and Australia would have to traverse relatively deep waters without coastal connections. Although few studies have reported habitat

utilization and movement patterns of adult bull sharks, Brunnschweiler *et al.* (2010) documented short range movements and fidelity to coastal areas by adults. Strong site fidelity to specific estuaries exhibited by juveniles has been documented by multiple authors (Yeiser *et al.* 2008; Heithaus *et al.* 2009; Ortega *et al.* 2009; Heupel *et al.* 2010; Heupel & Simpfendorfer 2011). Thus, I suspect that the habitat preferences of bull sharks restrict their migratory pathways to relatively shallow near shore areas with less frequent movements across relatively deep waters such as those separating Fiji from eastern Australia.

It is worth noting that although the sample size from the eastern Pacific is quite small (n=5), in the individual based STRUCTURE and PCOA analyses these samples cluster with samples from the western North Atlantic. One possible explanation for this apparent genetic similarity other than stochastic effects due to small sample sizes, is that bull shark migration has occurred between the western Atlantic and eastern Pacific through the Panama Canal. Notably, tarpon (*Megalops atlanticus*) and snook (*Centropomus undecimalis*) which are frequent prey items for bull sharks (Snelson & Williams 1981) have been reported in the Panama Canal (Hildebrand 1939; McCosker & Dawson 1975), and there are current recreational fisheries for both species in the Panama Canal. Bull sharks can tolerate very low salinities and have been reported to travel far upstream in freshwater rivers and lakes (Thorson 1971, 1972; Thorson 1976; Thomerson *et al.* 1977; Montoya & Thorson 1982), thus it is not inconceivable that some individuals have traversed the Panama Canal from the Atlantic into the eastern Pacific in search of prey or suitable habitat. This possibility could be investigated through a combination of tagging studies and genetic analyses of additional individuals from the eastern Pacific.

Population demographics and migration

The results of the demographic analyses are mixed. The M-ratio analyses indicated no change in population size. In contrast, BOTTLENECK analyses indicated population expansions in the Atlantic and Indo-Australia, with no change in population size in Fiji. However, the coalescent MSVAR analyses indicated a slight decline in the Atlantic with large increases in Fiji and Indo-Australia, while LAMARC was inconclusive and indicated possible small declines or small increases in all three populations. Additionally, Θ was >

1 for all populations based on the LAMARC analyses, thus the M-ratio test results are likely not indicative of a bottleneck. The effective population size estimates also vary between the analyses. The LAMARC-based estimated effective size for the Atlantic is similar to the effective population size reported by Karl *et al.* (2011), with estimates for Indo-Australia and Fiji in the same range. However the MSVAR estimates are unexpectedly approximately an order of magnitude larger than the LAMARC estimates. According to simulations comparing BOTTLENECK, M-RATIO and MSVAR (GIROD *ET AL.* 2011), MSVAR detected population size changes more reliably than either of the moment-based analyses. Additionally the authors reported that MSVAR scaled parameters are more accurate than natural parameters and both scaled and natural parameters are poorly estimated during population expansions. However, MSVAR was accurate in detecting increases in population size, and the authors recommend caution in interpreting the precise value of either set of parameters in the case of population expansion. Hence I interpret these results as meaning that the effective population sizes obtained from LAMARC are likely reasonable point estimates of the parameters and the MSVAR results are generally indicative of past demographic changes but not specific parameter estimates. Finally, the average microsatellite mutation rate calculated by MSVAR (3.57×10^{-6}) seems reasonable given that reported microsatellite mutation rates for fish are in the range of 10^{-4} to 10^{-5} (Shimoda *et al.* 1999; Yue *et al.* 2007), the reported range for mammals is 10^{-2} to 10^{-5} (Ellegren 2004), and mitochondrial and nuclear DNA appears to mutate approximately an order of magnitude more slowly in sharks than in other taxa (Martin *et al.* 1992; Martin 1999).

Although I found evidence of low levels of contemporary migration among the three genetically distinct populations, only the single migrant from the Atlantic into South Africa was supported by multiple analyses. Additionally, the directional estimates of historical migration rates were all fairly low. Thus, although migration among genetically distinct populations is possibly ongoing, it is probably not occurring at levels sufficient to rebuild populations or even to homogenize genetic diversity.

Conclusions and implications for management and conservation

The bull shark is an important component of recreational and commercial coastal fisheries. Although significant population declines have been suggested for many sharks due to overfishing (Baum & Blanchard 2010; Hisano *et al.* 2011), I find levels of genetic diversity in bull sharks that are among the highest for any elasmobranch species reported to date, suggesting that the genetic diversity of the species may be preserved through conservation programs designed to maintain relatively large effective population sizes. However, the species likely has a long generation time considering that the average female age at maturity is 13 - 18 years (Branstetter & Stiles 1987; Wintner *et al.* 2002; Fowler *et al.* 2005). Thus there may not have been sufficient time for contemporary population declines due to anthropogenic effects to impart their characteristic genetic signature.

Consistent with other published reports for bull sharks and despite the use of more microsatellite markers, I find no evidence of nuclear genetic structure along contiguous coastlines connecting very large distances. However, the clear nuclear differentiation between the Atlantic, IndoAustralia, and Fijian samples reported here, as well as the differentiation at mitochondrial loci over much smaller distances within oceanic regions reported by Karl *et al.* (2011) and Tillett *et al.* (2012) mean that effective management of these sharks must include local, regional, and international cooperation. Additionally, the previously reported female philopatry implies that subpopulations within oceanic basins are likely to be independent demographically over ecological timescales, and that further work to identify discrete matrilineal populations along continuous coastlines throughout the species range is imperative to formulating effective management policies.

Acknowledgements

I thank the following individuals for bull shark samples used in this study: G. Adkinson, D. Almojil, J. Apurdo, E. Brooks, J. Brunnschweiler, D. Chapman, G. Cliff, A. Danylchuck, B. Espinos, F. Galvan-Magana, S. Gulak, L. Hale, N. Hammerschlag, M. Heithaus, E. Hoffmayer, R. Jabado, C. Jones, T. Leary, P. Matich, R. McAuley, T. Neahr, L. Noble, L. Noll, N. Phillips, R. Pillans, W. Robbins, W. Smith, T. Testerman, R. Ward, J. Weary, J. Whitty, T. Wiley, C. Wilson, B. Winner, and S. Wintner. Financial support

for this project was received from Save Our Seas Foundation, Guy Harvey Ocean Foundation and Florida Sea Grant.

References

Au DW, Smith SE, Show C (2009) Shark productivity and reproductive protection, and a comparison with teleosts. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, EA Babcock). Blackwell Publishing Ltd., New Jersey.

Baum JK, Blanchard W (2010) Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research, 102, 229-239.

Baum JK, Myers RA, Kehler DG, Worm B, Harley SJ, Doherty PA (2003) Collapse and conservation of shark populations in northwest Atlantic. Science, 299, 389-392.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire génome, populations, interactions, CNRS UMR, 5000, 1996-2004.

Benavides M*, et al.* (2011) Global phylogeography of the dusky shark *Carcharhinus obscurus*: implications for fisheries management and monitoring the shark fin trade. Endangered Species Research, 14, 13-22.

Branstetter S, Burgess G (1997) Commercial shark fishery observer program. MARFIN Award NA57FF0286.

Branstetter S, Stiles R (1987) Age and growth estimates of the bull shark, *Carcharhinus leucas*, from the northern Gulf of Mexico. Environmental Biology of Fishes, 20, 169-181.

Brunnschweiler JM, Baensch H (2011) Seasonal and Long-Term Changes in Relative Abundance of Bull Sharks from a Tourist Shark Feeding Site in Fiji. PLOS One, 6, e16597.

Brunnschweiler JM, Queiroz N, Sims DW (2010) Oceans apart? Short-term movements and behaviour of adult bull sharks *Carcharhinus leucas* in Atlantic and Pacific Oceans determined from pop-off satellite archival tagging. Journal of Fish Biology, 77, 1343-1358.

Carlson JK, Ribera MM, Conrath CL, Heupel MR, Burgess GH (2010) Habitat use and movement patterns of bull sharks *Carcharhinus leucas* determined using pop-up satellite archival tags. Journal of Fish Biology, 77, 661-675.

Casey JG, Hoey JJ (1985) Estimated catches of large sharks by US recreational fishermen in the Atlantic and Gulf of Mexico. Shark catches from selected fisheries off the US East Coast. NOAA Technical Report, NMFS SSRF.

Castro JI (1983) The Sharks of North American Waters Texas A. & M. University Press, College Station, USA.

Coombs JA, Letcher BH, Nislow KH (2008) CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. Molecular Ecology Resources, 8, 578-580.

Cortes E (2002) Incorporating Uncertainty into Demographic Modeling: Application to Shark Populations and Their Conservation. Conservation Biology, 16, 1048-1062.

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. Molecular Ecology Resources, 10, 556-557.

Cunha RL, Coscia I, Madeira C, Mariani S, Stefanni S, Castilho R (2012) Ancient Divergence in the Trans-Oceanic Deep-Sea Shark *Centroscymnus crepidater*. PLOS One, 7, e49196.

Curtis TH, Adams DH, Burgess GH (2011) Seasonal distribution and habitat associations of bull sharks in the Indian River Lagoon, Florida: A 30-Year synthesis. Transactions of the American Fisheries Society, 140, 1213-1226.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW (2012) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). PLOS One, 7, e29986.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 1-38.

Duchesne P, Turgeon J (2009) FLOCK: a method for quick mapping of admixture without source samples. Molecular Ecology Resources, 9, 1333-1344.

Duchesne P, Turgeon J (2012) FLOCK Provides Reliable Solutions to the "Number of Populations" Problem. Journal of heredity, 103, 734-743.

Dulvy ND, Forrest RE (2010) Life histories, population dynamics, and extinction risks in chondrichthyans. In: Sharks and their relatives II: Biodiversity, adaptive physiology and conservation (eds. JC Carrier, JA Musick, MR Heithaus), pp. 635-676. CRC Press, Boca Raton, FL.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359-361.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. Theoretical applications in genetics, 92, 832-839.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics, 5, 435-445.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14, 2611-2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics, 164, 1567-1587.

Feldheim KA, Gruber SH, Ashley MV (2002) The breeding biology of lemon sharks at a tropical nursery lagoon. Proceedings of the Royal Society B, 269, 1655-1661.

Fowler SL, Cavanagh RD, Camhi M, Burgess GH, Cailliet GM, Fordham SV, Simpfendorfer CA, Musick JA (2005) Sharks, rays and chimeras: the status of the Chondrichthyan fishes, pp. 1-476. IUCN.

Froeschke JT, Froeschke BF, Stinson CM (2012) Long-term trends of bull shark (*Carcharhinus leucas*) in estuarine waters of Texas, USA. Canadian Journal of Fisheries and Aquatic Sciences, 70, 13-21.

Garcia VB, Lucifora LO, Myers RA (2008) The importance of habitat and life history to extinction risk in sharks, skates, rays and chimaeras. Proceedings of the Royal Society B, 275, 83-89.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology, 10, 305-318.

Girod C, Vitalis R, Leblois R, Freville H (2011) Inferring Population Decline and Expansion From Microsatellite Data: A Simulation-Based Evaluation of the Msvar Method. Genetics, 188, 165-U287.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of heredity, 86, 485-486.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Guo SW, Thompson EA (1992) Performing exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics, 48, 361-372.

Heithaus MR, Burkholder D, Hueter RE, Heithaus LI, Pratt HL, Jr., Carrier JC (2007) Spatial and temporal variation in shark communities of the lower Florida keys and evidence for historical population declines. Canadian Journal of Fisheries and Aquatic Sciences, 64, 1302+.

Heithaus MR, Delius BK, Wirsing AJ, Dunphy-Daly MM (2009) Physical factors influencing the distribution of a top predator in a subtropical oligotrophic estuary. Limnology and Oceanography, 54, 472-482.

Heupel MR, Carlson JK, Simpfendorfer CA (2007) Shark nursery areas: concepts, definition, characterization and assumptions. Marine Ecology-Progress Series, 337, 287-297.

Heupel MR, Simpfendorfer CA (2011) Estuarine nursery areas provide a low-mortality environment for young bull sharks *Carcharhinus leucas*. Marine Ecology Progress Series, 433, 237-244.

Heupel MR, Yeiser BG, Collins AB, Ortega L, Simpfendorfer CA (2010) Long-term presence and movement patterns of juvenile bull sharks, *Carcharhinus leucas*, in an estuarine river system. Marine and Freshwater Research, 61, 1-10.

Hildebrand SF (1939) The Panama Canal as a passageway for fishes, with list and remarks on the fishes and invertebrates. Zoologica, New York, 24, 15-45.

Hisano M, Connolly SR, Robbins WD (2011) Population growth rates of reef sharks with and without fishing on the Great Barrier Reef: Robust estimation with multiple models. PLOS One, 6, e25028.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources, 9, 1322-1332.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, 23, 1801-1806.

Kalinowski ST, Wagner AP, Taper ML (2006) ml-relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes, 6, 576-579.

Karl SA, Castro ALF, Lopez JA, Charvet P, Burgess GH (2011) Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. Conservation Genetics, 12, 371-382.

Keeney D, Heupel M, Hueter R, Heist E (2005) Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. Molecular Ecology, 14, 1911-1923.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics, 22, 768-770.

Martin A, Naylor G, Palumbi S (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature, 357, 1992.

Martin AP (1999) Substitution rates of organelle and nuclear genes in sharks: implicating metabolic rate (again). Molecular Biology and Evolution, 16, 996-1002.

Matich P, Heithaus M (2012) Effects of an extreme temperature event on the behavior and age structure of an estuarine top predator, *Carcharhinus leucas*. Marine Ecology Progress Series, 447, 165-178.

Matich P, Heithaus MR (2014) Multi-tissue stable isotope analysis and acoustic telemetry reveal seasonal variability in the trophic interactions of juvenile bull sharks in a coastal estuary. Journal of Animal Ecology, 83, 199-213.

McCosker JE, Dawson CE (1975) Biotic passage through the Panama Canal, with particular reference to fishes. Marine Biology, 30, 343-351.

Montoya R, Thorson T (1982) The bull shark (*Carcharhinus leucas*) and largetooth sash (*Pristis perotteti*) in Lake Bayano, a tropical man-made impoundment in Panama. Environmental Biology of Fishes, 7, 341-347.

Myers RA, Baum JK, Shepherd TD, Powers SP, Peterson CH (2007) Cascading effects of the loss of apex predatory sharks from a coastal ocean. Science, 315, 1846-1850.

Nei M (1987) Molecular evolutionary genetics Columbia University Press, New York.

O'Connell MT, Shepherd TD, O'Connell AM, Myers RA (2007) Long-term declines in two apex predators, bull sharks (*Carcharhinus leucas*) and alligator gar (*Atractosteus spatula*), in Lake Pontchartrain, an oligohaline estuary in southeastern Louisiana. Estuaries and Coasts, 30, 567-574.

Ortega LA, Heupel MR, Van Beynen P, Motta PJ (2009) Movement patterns and water quality preferences of juvenile bull sharks (*Carcharhinus leucas*) in a Florida estuary. Environmental Biology of Fishes, 84, 361-373.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology, 13, 55-65.

Pardini A*, et al.* (2001) Sex-biased dispersal of great white sharks. Nature, 412, 139-140.

Park S (2001) Trypanotolerance in West African cattle and the population genetic effects of selection, PhD Dissertation, University of Dublin, Dublin, Ireland.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of heredity, 95, 536-539.

Piry S, Luikart G, Cornuet J-M (1999) BOTILENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of heredity, 90, 502-503.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA, 94, 9197-9201.

Rosenberg NA (2004) distruct: a program for the graphical display of population structure. Molecular Ecology Notes, 4, 137-138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8, 103-106.

Rousset F, Raymond M (1995) Testing heterozygote excess and deficiency. Genetics, 140, 1413-1419.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18, 233-234.

Schultz JK, Feldheim KA, Gruber SH, Ashley MV, McGovern TM, Bowen BW (2008) Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). Molecular Ecology, 17, 5336-5348.

Shepherd TD, Myers RA (2005) Direct and indirect fishery effects on small coastal elasmobranchs in the northern Gulf of Mexico. Ecology Letters, 8, 1095-1104.

Shimoda N*, et al.* (1999) Zebrafish Genetic Map with 2000 Microsatellite Markers. Genomics, 58, 219-232.

Simpfendorfer C, Milward N (1993) Utilization of a tropical bay as a nursery area by sharks of the families Carcharhinidae and Sphyrnidae. Environmental Biology of Fishes, 37, 337-345.

Snelson FF, Mulligan TJ, Williams SE (1984) Food Habits, Occurrence, and Population Structure of the Bull Shark, *Carcharhinus Leucas*, in Florida Coastal Lagoons. Bulletin of Marine Science, 34, 71-80.

Snelson FFJ, Williams SE (1981) Notes on the occurrence, distribution, and biology of elasmobranch fishes in the Indian River Lagoon system, Florida. Estuaries, 4, 110-120.

Speed CW, Field IC, Meekan MG, Bradshaw CJA (2010) Complexities of coastal shark movements and their implications for management. Marine Ecology Progress Series, 408, 275-293.

Stevens J, Bonfil R, Dulvy N, Walker P (2000) The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. ICES Journal of Marine Science, 57, 476-494.

Storz JF, Beaumont MA (2002) Testing for genetic evidence of population expansion and contraction: An empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. Evolution, 56, 154-166.

Taguchi M, Kitamura T, Yokawa K (2011) Genetic population structure of shortfin mako (*Isurus oxyrinchus*) inferred from mitochondrial DNA on interoceanic scale. ISC/11/SHARKWG-1/02.

Testerman CB, Fitzpatrick S, Prodohl P, Clarke S, Chapman D, Simpfendorfer C, Shivji MS (In prep-a) Contrasting patterns of genetic structure in the great hammerhead shark (*Sphyrna mokarran*) reveal a third example of active anti-Agulhas dispersal.

Testerman CB, Shivji MS, al. e (in prep-b) Global phylogeography of the amphi-temperate smooth hammerhead shark *Sphyrna zygaena* indicates female philopatry and dispersal around Cape Horn.

Thomerson JE, Thorson TB, Hempel RL (1977) The Bull Shark, *Carcharhinus leucas*, from the Upper Mississippi River near Alton, Illinois. Copeia, 1977, 166-168.

Thorson TB (1971) Movements of bull sharks, *Carcharhinus leucas*, between the Caribbean Sea and Lake Nicaragua demonstrated by tagging. Copeia, 1971, 336-338.

Thorson TB (1972) The Status of the Bull Shark, *Carcharhinus leucas*, in the Amazon River. Copeia, 1972, 601-605.

Thorson TB (1976) The status of the Lake Nicaragua shark: an updated appraisal.

Tillett BJ, Meekan MG, Field IC, Thorburn DC, Ovenden JR (2012) Evidence for reproductive philopatry in the bull shark *Carcharhinus leucas*. Journal of Fish Biology, 80, 2140-2158.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535-538.

Watson RA, Cheung WWL, Anticamara JA, Sumaila RU, Zeller D, Pauly D (2013) Global marine yield halved as fishing intensity redoubles. Fish and Fisheries, 14, 493-503.

Wiley TR, Simpfendorfer CA (2007) The ecology of elasmobranchs occurring in the Everglades National Park, Florida: implications for conservation and management. Bulletin of Marine Science, 80, 171-189.

Wintner SP, Dudley SFJ, Kistnasamy N, Everett B (2002) Age and growth estimates for the Zambezi shark, *Carcharhinus leucas*, from the east coast of South Africa. Marine and Freshwater Research, 53, 557-566.

Yeiser BG, Heupel MR, Simpfendorfer CA (2008) Occurrence, home range and movement patterns of juvenile bull (*Carcharhinus leucas*) and lemon (*Negaprion brevirostris*) sharks within a Florida estuary. Marine and Freshwater Research, 59, 489-501.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). Genetica, 129, 329-331.

Tables

Table 1. *Carcharhinus leucas* summary statistics for each microsatellite locus by collection area. *n* = number of individuals, *Na* = number of alleles, A_r = allelic richness, GD = genetic diversity (Nei, 1987), H_E = expected frequency of heterozygotes, H_O = observed frequency of heterozygotes, Null = frequency of null alleles estimated using the Expectation Maximization (EM) algorithm of Dempster, Laird, and Rubin (1977). Far right column contains for each locus across all populations: (i) mean H_E, H_O values, (ii) overall A_r and F_{IS}, and (iii) total *Na* and null allele frequency. The bottom row contains for each population across all loci: (i) mean number of alleles, (ii) observed and expected frequency of heterozygotes, (iii) F_{IS}, (iv) mean null allele frequency. Last two lines of the bottom row, far right column are the overall F_{IS} and mean null allele frequency for all loci in all populations. Bold F_{IS} values are significant after sequential Bonferroni correction ($P < 0.05$), bold null allele values are > 0.10. Samples from GoC (*n*= 5) and PH (*n* = 1) not included due to low sample sizes. Collection areas are coded as in Figure 1.

Locus	WNA *n* = 30	FLK *n* = 14	Evg *n* = 95	GoM *n* = 63	SAfr *n* = 30	MEAS *n* = 23	MEPG *n* = 18	IndoAnd *n* = 44	WAus *n* = 17	EAus *n* = 58	Fiji *n* = 70	Mean/locus Total/locus
CL07												
Na	23	15	32	32	18	22	16	23	16	26	19	41
A_r	14.971	15	15.485	16.465	10.671	15.303	12.646	11.834	13.285	12.18	11.232	15.111
GD	0.951	0.958	0.953	0.964	0.849	0.945	0.913	0.887	0.93	0.899	0.902	0.9430
H_e	0.9333	0.9132	0.9475	0.9551	0.8341	0.9225	0.8873	0.8742	0.8997	0.8901	0.8956	0.9415
H_o	0.8333	0.8333	0.9121	0.9016	0.8621	0.9091	0.8889	0.7561	0.8235	0.8182	0.8714	0.8632
F_{IS}	0.1239	0.1304	0.0429	0.0642	-0.016	0.0378	0.0268	0.1472	0.1146	0.0899	0.0342	**0.0843**
Null	0.0496	0.0158	0.0223	0.0232	0.0000	0.0303	0.0000	0.0599	0.0381	0.0279	0.0138	0.0255
CL102												
Na	2	1	1	2	5	5	6	6	7	5	4	9
A_r	1.414	1	1	1.194	4.246	4.319	5.294	3.947	5.529	4.437	3.857	3.776
GD	0.034	0	0	0.016	0.675	0.627	0.696	0.598	0.619	0.665	0.704	0.4890
H_e	0.0339	0	0	0.016	0.6633	0.6125	0.679	0.5886	0.5969	0.6578	0.6987	0.4884
H_o	0.0345	0	0	0.0161	0.6786	0.5652	0.7778	0.3864	0.4706	0.5	0.6857	0.3289
F_{IS}	0	0	0	0	-0.0049	0.0992	-0.1174	**0.3537**	0.2404	0.2481	0.0257	**0.3274**
Null	0.0000	0.0010	0.0010	0.0000	0.0097	0.0000	0.0000	**0.1404**	0.0396	0.0904	0.0186	0.0273
CL107												
Na	16	13	22	24	20	17	14	21	17	24	17	37
A_r	10.243	11.81	9.846	9.836	14.631	13.068	12.058	13.888	14.273	13.574	11.508	14.28
GD	0.851	0.885	0.842	0.831	0.955	0.935	0.904	0.948	0.95	0.942	0.923	0.941
H_e	0.8372	0.8495	0.8368	0.8234	0.9375	0.9132	0.877	0.9357	0.9221	0.9337	0.9163	0.9401
H_o	0.8667	0.7857	0.734	0.6984	0.9286	0.9091	0.9375	0.8182	0.9412	0.9643	0.9571	0.8458
F_{IS}	-0.0182	0.1118	**0.1281**	0.1596	0.0277	0.0278	-0.0369	**0.1369**	0.0097	-0.0238	-0.0374	**0.1013**
Null	0.0000	0.0000	0.0454	0.0700	0.0068	0.0224	0.0000	0.0589	0.0000	0.0000	0.0000	0.0185
CL108												
Na	22	14	32	30	21	23	17	29	19	31	24	40
A_r	13.533	13.077	13.947	14.476	14.558	16.936	14.358	15.916	16.101	16.005	13.773	15.903
GD	0.932	0.94	0.927	0.939	0.952	0.97	0.947	0.961	0.963	0.96	0.942	0.9590
H_e	0.9167	0.9082	0.9222	0.9304	0.9343	0.949	0.9187	0.9496	0.9316	0.9518	0.9351	0.9575
H_o	0.9333	1	0.8913	0.9	0.9286	0.9565	0.9412	0.9091	0.9375	0.9483	0.8571	0.9115
F_{IS}	-0.0012	-0.0643	0.039	0.0411	0.0243	0.0143	0.0058	0.0542	0.026	0.0124	0.0905	**0.0491**
Null	0.0090	0.0000	0.0125	0.0000	0.0029	0.0000	0.0000	0.0134	0.0000	0.0190	0.0358	0.0084

Table 1 (Continued)

Locus	WNA *n* = 30	FLK *n* = 14	Evg *n* = 95	GoM *n* = 63	SAfr *n* = 30	MEAS *n* = 23	MEPG *n* = 18	IndoAnd *n* = 44	WAus *n* = 17	EAus *n* = 58	Fiji *n* = 70	Mean/locus Total/locus
CPL90												
Na	23	13	25	21	26	22	22	26	17	29	22	38
A_r	14.29	13	13.295	12.426	16.234	15.17	16.82	15.703	14.107	15.57	13.187	15.072
GD	0.939	0.909	0.934	0.923	0.961	0.952	0.962	0.96	0.945	0.958	0.933	0.9530
H_e	0.9233	0.875	0.9284	0.9153	0.9461	0.9319	0.9367	0.9486	0.9187	0.9494	0.926	0.9515
H_o	0.9333	1	0.913	0.8889	1	1	1	0.9535	1	0.9298	0.9571	0.9429
F_{IS}	0.0061	-0.1	0.022	0.0369	-0.04	-0.0509	-0.039	0.0066	-0.0584	0.0294	-0.0264	0.0102
Null	0.0000	0.0000	0.0138	0.0012	0.0000	0.0000	0.0000	0.0055	0.0000	0.0011	0.0000	0.0020
CS02												
Na	9	8	15	11	16	17	15	21	17	19	18	37
A_r	5.768	7.538	6.305	5.236	9.622	11.006	13.806	11.918	14.047	10.458	10.991	11.28
GD	0.64	0.769	0.671	0.645	0.716	0.804	0.945	0.873	0.954	0.797	0.889	0.8810
H_e	0.6288	0.7449	0.6674	0.6399	0.6939	0.7807	0.9082	0.8554	0.9135	0.7848	0.8763	0.8800
H_o	0.6429	0.8571	0.7204	0.6066	0.1786	0.5217	0.8571	0.359	0.5294	0.25	0.2063	0.4885
F_{IS}	-0.0041	-0.1143	-0.0741	0.0603	**0.7507**	**0.3514**	0.093	**0.5889**	**0.4451**	**0.6862**	**0.7678**	**0.4458**
Null	0.0000	0.0000	0.0063	0.0324	**0.3086**	**0.1510**	0.0376	**0.2717**	**0.2288**	**0.3077**	**0.3574**	0.1547
CS07												
Na	6	4	7	4	15	17	14	17	10	17	11	28
A_r	4.755	4	4.546	3.908	11.402	12.54	11.813	9.727	10	9.757	7.266	10.701
GD	0.722	0.747	0.707	0.679	0.909	0.916	0.907	0.876	0.924	0.887	0.796	0.9010
H_e	0.71	0.7168	0.7025	0.672	0.8848	0.8905	0.8735	0.8652	0.8646	0.8771	0.7895	0.8993
H_o	0.7333	0.6429	0.6559	0.5574	0.5185	0.7273	0.6111	0.8409	0.4167	0.6481	0.6866	0.6561
F_{IS}	-0.0159	0.1397	0.0717	0.1786	**0.4295**	0.2057	**0.3261**	0.0395	**0.5492**	**0.2697**	0.1377	**0.2715**
Null	0.0000	0.0152	0.0234	0.0661	**0.2039**	0.0825	**0.1447**	0.0297	**0.2393**	**0.1321**	0.0600	0.0906
CT05												
Na	7	4	11	7	9	8	9	10	11	11	8	15
A_r	5.942	4	6.321	5.286	7.584	6.918	7.703	7.209	9.329	7.374	4.15	7.569
GD	0.788	0.772	0.816	0.783	0.767	0.751	0.77	0.747	0.842	0.781	0.62	0.8380
H_e	0.77	0.727	0.8098	0.7731	0.7503	0.7353	0.7405	0.7368	0.8166	0.7736	0.6146	0.8366
H_o	0.5	0.2857	0.5474	0.4	0.5862	0.7826	0.5294	0.6829	0.8235	0.6909	0.5362	0.5689
F_{IS}	**0.3654**	**0.6299**	**0.3288**	**0.489**	0.2353	-0.0421	**0.3126**	0.0853	0.0218	0.1159	0.1347	**0.3210**
Null	**0.1470**	**0.2570**	**0.1495**	**0.2128**	0.0679	0.0033	**0.1074**	0.0181	0.0000	0.0565	0.0650	0.0986
LS24												
Na	5	3	4	5	5	5	7	5	5	6	6	8
A_r	3.965	2.999	3.23	3.655	4.111	4.667	6.093	4.101	4.737	4.35	4.77	4.501
GD	0.616	0.61	0.576	0.603	0.63	0.7	0.696	0.662	0.708	0.631	0.659	0.6360
H_e	0.6034	0.5842	0.5726	0.5993	0.6183	0.6853	0.6728	0.6552	0.6914	0.6241	0.6542	0.6356
H_o	0.4828	0.5	0.5158	0.6825	0.5862	0.7391	0.5556	0.75	0.875	0.5172	0.6	0.6013
F_{IS}	0.2168	0.1802	0.1044	-0.1311	0.0694	-0.0565	0.2019	-0.1334	-0.2353	0.1797	0.09	0.0551
Null	0.0522	0.0847	0.0154	0.0000	0.0000	0.0000	**0.1190**	0.0000	0.0000	0.0766	0.0692	0.0379
PG11												
Na	5	3	5	6	9	6	8	6	5	10	5	13
A_r	3.649	2.714	2.311	3.034	5.924	4.819	6.561	4.186	4.545	5.345	3.721	5.478
GD	0.326	0.14	0.16	0.308	0.685	0.708	0.748	0.647	0.597	0.69	0.514	0.7370
H_e	0.3206	0.1352	0.1587	0.3059	0.6733	0.69	0.7299	0.6402	0.5761	0.6842	0.5102	0.7361
H_o	0.3333	0.1429	0.1702	0.3226	0.6667	0.5652	0.8333	0.7045	0.4706	0.7193	0.5429	0.4662
F_{IS}	-0.0229	-0.0196	-0.067	-0.0463	0.0268	0.2022	-0.1135	-0.0891	0.2123	-0.0425	-0.0568	**0.3676**
Null	0.0000	0.0000	0.0000	0.0265	0.0373	0.0900	0.0000	0.0000	0.0867	0.0256	0.0000	0.0242
PG13												
Na	9	8	13	15	20	18	17	22	16	21	18	31
A_r	6.793	7.665	6.634	7.656	14.291	13.676	13.311	13.722	13.615	11.778	11.035	12.284
GD	0.705	0.788	0.746	0.772	0.952	0.944	0.922	0.942	0.947	0.917	0.908	0.9120
H_e	0.6922	0.7551	0.742	0.7659	0.9352	0.9234	0.8951	0.9313	0.917	0.9084	0.9015	0.9114
H_o	0.6667	0.6429	0.7579	0.7541	0.931	0.9565	0.8889	0.9773	0.8824	0.8909	0.8857	0.8355
F_{IS}	0.0538	0.1847	-0.0161	0.0237	0.022	-0.0136	0.0355	-0.0379	0.068	0.0285	0.0247	**0.0843**
Null	0.0220	0.0731	0.0000	0.0266	0.0028	0.0000	0.0000	0.0000	0.0048	0.0039	0.0064	0.0127
All loci												
Na	12	8	15	14	15	15	13	17	13	18	14	27
H_e	0.67	0.6554	0.6625	0.6724	0.8065	0.8213	0.829	0.8164	0.8226	0.8214	0.7925	0.8344
H_o	0.6327	0.6082	0.6198	0.6117	0.715	0.7848	0.8019	0.7398	0.7428	0.7161	0.7078	0.6826
F_{IS}	0.0726	0.1101	0.0698	0.0985	0.1310	0.0670	0.0627	0.1056	0.1290	0.1370	0.1141	**0.1829**
Null	0.0254	0.0406	0.0263	0.0417	0.0582	0.0345	0.0372	0.0543	0.0579	0.0673	0.0569	0.0455

Table 2. *C. leucas* pairwise F_{ST} and Jost D_{est}

	WNA	IndoAus	Fiji
WNA		0.4652	0.5262
IndoAus	0.1582		0.0839
Fiji	0.1995	0.0284	

F_{ST} below the axis and Jost D_{est} above

F_{ST} P all significant ($P < 0.05$)

Table 3. *C. leucas* microsatellite demographic analyses.

Table 3a. BOTTLENECK: Significance was estimated using the Wilcoxon sign rank deficiency test as $P < 0.05$. MRATIO: M-values were estimated by M_P_val and the critical values were estimated using Critical_M with a range of pre-bottleneck θ (0.01, 0.1, 1, 10). MSVAR: reported values include the current population size (N0), the historic population size (N1), the time in million years since population contraction or expansion, and their 95% highest posterior distribution (HPD).

	Bottleneck		Mratio									
	Heterozygote		$M < M_c$	MSvar Results N_0			MSvar Results N_1			MSvar Time since event		
	Excess	Deficit	When:	mode	95% HPD		mode	95% HPD		mode	95% HPD	
Population												
Atl	ns	significant	θ <= 1	2.E+06	1.E+01	3.E+08	9.E+06	7.E+03	2.E+10	5.E+06	4.E+03	1.E+10
IndoAus	ns	significant	θ <= 0.1	1.E+07	2.E+02	6.E+08	3.E+05	7.E+03	2.E+10	4.E+06	2.E+03	6.E+09
Fiji	ns	ns	θ <= 0.1	7.E+06	1.E+02	4.E+08	2.E+05	5.E+03	2.E+10	3.E+06	1.E+03	6.E+09

Table 3b. LAMARC: Reported values are the effective population size of each population (N_e), exponential growth rate parameter (Population growth), and migrants per generation with m1➔2 indicating migrants from the second population into the first, and vice versa.

	Lamarc Results N_e			Lamarc Results	Lamarc Results			Lamarc Results		
	MPE	5% CI	95% CI	Population growth	# migr per gen 1 ➔ 2			# migr per gen 2 ➔ 1		
Population					MPE	5% CI	95% CI	MPE	5% CI	95% CI
Atl	350,216	350,142	350,800	Spans 0						
IndoAus	425,028	276,319	638,701	Spans 0						
Fiji	173,854	107,437	221,203	Slightly positive						
Popn Comparison										
Atl-IndoAus					1.156	1.122	17.836	1.288	1.140	2.090
Atl-Fiji					0.563	0.401	0.591	0.793	0.430	1.095
IndoAus-Fiji					2.091	1.522	2.560	1.033	0.892	1.690

Table 4. *C. leucas* migrants. Values in parenthesis are posterior probabilities inferred from STRUCTURE analyses.

		Migrants		
Specimen	Collection area	1st generation GENECLASS	Capture location (probability of origin) STRUCTURE	Inferred origin (probability) STRUCTURE
OC-039	SAfr		Indo-Aus (0.32)	Fiji (0.56)
OC-040	SAfr		Indo-Aus (0.22)	Fiji (0.63)
OC-041	SAfr		Indo-Aus (0.05)	Fiji (0.89)
OC-111	SAfr	Migr from WNA	Indo-Aus (0.00)	WNA (1.00)
OC-044	Fiji	Migr from IndoAus	Fiji (0.90)	Fiji (0.90)
OC-361	Fiji	Migr from IndoAus	Fiji (0.95)	Fiji (0.95)
OC-507	Fiji	Migr from IndoAus	Fiji (0.94)	Fiji (0.94)
OC-565	Fiji	Migr from IndoAus	Fiji (0.70)	Fiji (0.70)

Figures

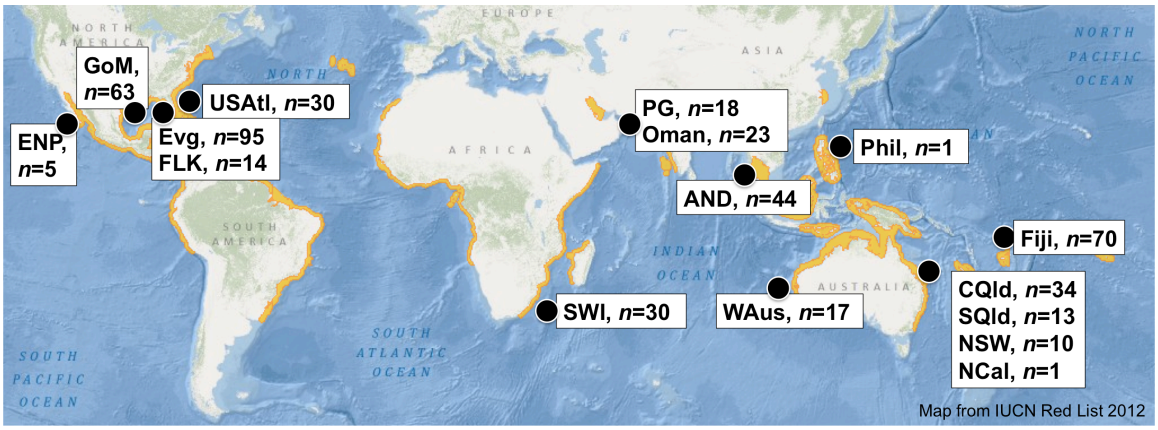


Fig. 1. Map of *C. leucas* distribution and sample sizes. Known global distribution shown in gold. General collection areas are represented by a black circle and sample sizes are indicated in boxes. Collection localities are coded as follows: (1) western North Atlantic Ocean: US Atlantic (USAtl); Florida Keys (FLK); Florida Everglades (Evg); Gulf of Mexico (GoM); (2) Indian Ocean: western Indian Ocean (SAfr – South Africa); Arabian Sea (MEAS); Persian Gulf (MEPG); Andaman Sea (IndoAnd); eastern Indian Ocean (WAus – western Australia); and (3) Pacific Ocean: eastern Australia (EAus) – Central Queensland (CQld), South Queensland (SQld), New South Wales (NSW), New Caledonia (NCal); Fiji; Philippines (PH); and eastern Pacific (EPac).

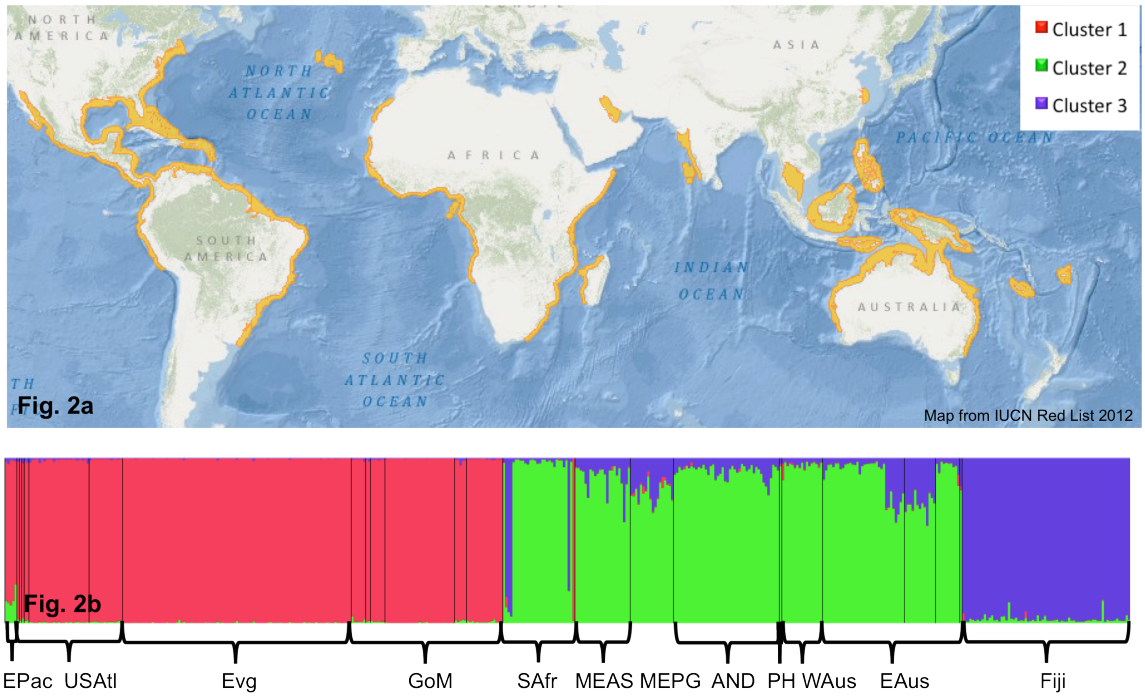


Fig 2. *C. leucas* microsatellite STRUCTURE analysis results. Fig. 2a: pie charts indicate the average proportional membership coefficient of individual sharks in the three distinct clusters inferred from nuclear microsatellite genotypes by the program STRUCTURE. Pie chart sizes are roughly proportional to sample sizes. Fig. 2b: assignment of individual sharks to each cluster in a conventional bar plot. Collection localities are coded as in Figure 1.

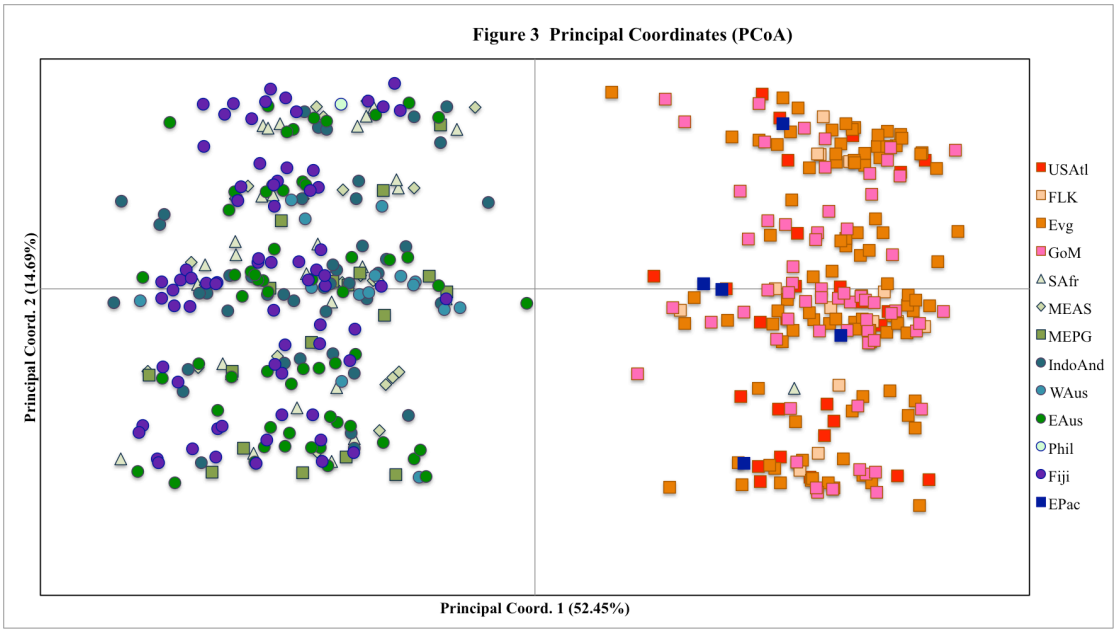


Fig 3. *C. leucas* Principal Coordinate Analysis (PCoA). Individuals from all sampling locations were included, small colored shapes indicate the locations according to the figure legends. The first two principal coordinate axes are shown with the amount of variance explained by each in parentheses. Collection areas are coded as in Figure 1.

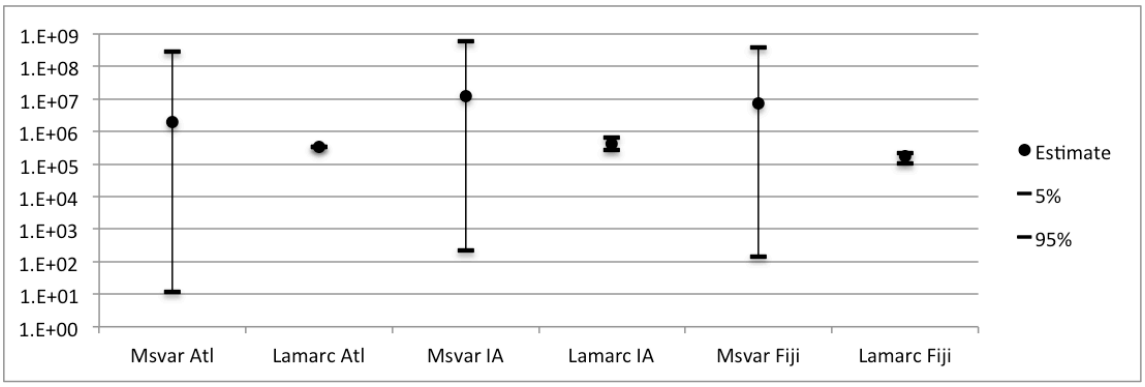


Fig. 4. Summary of the estimated effective population sizes (N_e) for *C. leucas* based on Bayesian inferences of microsatellite data. The x-axis indicates the populations (western North Atlantic - Atl, Indo-Australia - IA, or Fiji), and the analytical program used (MSVAR or LAMARC). The y-axis is the effective population size (N_e) in logarithmic scale, the central dot represents the point estimate (MSVAR - mode, LAMARC – most probable estimate (MPE)), the vertical lines extend from the 5% to the 95% highest posterior density (HPD) for MSVAR and percentiles for LAMARC.

Supplementary Information

Table S1. Microsatellite loci used to genotype *C. leucas*

Locus	Annealing temperature (°C)	Reference
Clim007	55	Keeney & Heist (2003)
Clim102	55	Keeney & Heist (2003)
Clim107	55	Keeney & Heist (2003)
Clim108	58	Keeney & Heist (2003)
Pg011	58	Fitzpatrick *et al.* (2011)
Pg013	55	Fitzpatrick *et al.* (2011)
Csor002	50	Ovenden *et al.* (2005)
Csor007	55	Ovenden *et al.* (2005)
Ctil005	50	Ovenden *et al.* (2005)
Ls024	58	Feldheim *et al.* (2001)
Cpl090	58	Portnoy *et al.* (2006)

CHAPTER 3: Global patterns of population genetic structure and population dynamics of the endangered great hammerhead shark (*Sphyrna mokarran*) revealed by mitochondrial and nuclear genetic analyses.

Abstract

The population status and dynamics of the great hammerhead shark (*Sphyrna mokarran*) are of considerable conservation interest due to its IUCN Endangered listing, extreme fishing bycatch mortality and high market value in the international shark fin market. Information on genetic structure and diversity and dispersal patterns can be used to inform conservation and management decisions. I evaluated the population genetic structure of *S. mokarran* based on globally distributed samples obtained from the western North Atlantic (US Atlantic, Gulf of Mexico and Caribbean), western South Atlantic, Indian Ocean (Red Sea, Persian Gulf, Indonesia, and western Australia) and south western Pacific (eastern Australia) using complete mitochondrial control region sequences (n=272, 1 082 nucleotides) and 12 nuclear microsatellite loci (n=264). The population structure results based on mitochondrial and microsatellite markers were generally concordant, revealing strong geographic partitioning between samples from the western Atlantic and Australia. I found no population structure within the western Atlantic or between east and west Australia. There is shallow but statistically significant genetic structuring among samples from the western Atlantic, northern Indian Ocean and Australia, albeit with some differences between the two marker types in fine scale patterns involving Red Sea. These differences are likely a result of past dispersal from the Atlantic into the Indian Ocean, combined with more recent male mediated gene flow between the northern Indian Ocean and Australia providing an example of active anti-Agulhas dispersal. Geographic patterns of genetic structure in this species suggests that genetic data may prove useful to source the broad geographic origin of *S. mokarran* fins in trade and assist with delineation of stocks for individual population assessment and general conservation efforts.

Introduction

It had been thought that marine species are typically panmictic throughout their range (Jennings *et al.* 2001) but many studies have revealed statistically significant genetic structuring over short and long distances (Knutsen *et al.* 2011). Recent studies examining phylogeography and dispersal patterns in broadly distributed, mobile marine species without planktonic phases (Benavides *et al.* 2011; Cunha *et al.* 2012; Dutton *et al.* 2013; Richards *et al.* 2013; Ruegg *et al.* 2013) have revealed complex patterns of genetic connectivity that reflect individual species biology as well as ecosystem and habitat usage. Knowledge of a species' overall population genetic structure and dispersal patterns can have important implications for management and conservation strategies.

Contemporary and historical processes can influence a species' population genetic structure and dispersal patterns (Wares 2002; Patarnello *et al.* 2007; Perez-Losada *et al.* 2007; Roy *et al.* 2007; Borrero-Perez *et al.* 2011). Determining the relative contribution of these events can be confounded by differences in the time required for lineage sorting in mitochondrial versus nuclear DNA (Palumbi *et al.* 2001). Indeed, discordance in population structure patterns revealed from nuclear and mitochondrial markers have been reported for several species (Welch *et al.* 2011; Bastos-Silveira *et al.* 2012; Singhal & Moritz 2012; Takahashi & Hori 2012; Warren *et al.* 2012); however, genetic methods based on coalescent theory can be particularly useful in distinguishing contemporary patterns of migration or sex-biased dispersal from patterns due to historical events and incomplete lineage sorting (Avise *et al.* 1983; Degnan & Rosenberg 2009).

Elasmobranchs are top predators in many marine ecosystems and are particularly vulnerable to human overexploitation due to life history traits such as late maturity, slow growth, long gestation periods, and low fecundity, resulting in elevated sensitivity to overharvesting (Cortes 2002; Garcia *et al.* 2008; Au *et al.* 2009; Dulvy & Forrest 2010). Severe declines in population sizes have been reported for many shark species (Baum & Myers 2004; Baum *et al.* 2005; Baum & Blanchard 2010) and removal of these predators may cause significant ecosystem effects and community restructuring through top-down predator release (Shepherd & Myers 2005; Myers *et al.* 2007; Heithaus *et al.* 2008; Ferretti *et al.* 2010). In general, despite their ecological and commercial importance little

is known about the population dynamics of most shark species beyond basic biological parameters.

The great hammerhead shark (*Sphyrna mokarran*) is the largest of the nine species in the family Sphyrnidae. It is circumglobally distributed, found primarily in coastal and continental shelf waters of warm temperate and tropical oceans (Compagno 1984; Last & Stevens 2009). The great hammerhead is not typically targeted by commercial fisheries in the Atlantic; however, there is cause for significant management and conservation concern as it is a frequent component of commercial bycatch and has on-line mortality rates greater than 90% (Morgan & Burgess 2007). Additionally, the very high market value of its large fins (Rose 1996; Abercrombie *et al.* 2005; Clarke *et al.* 2006) provides incentive for increased opportunistic fishing efforts. Although species specific population trends are rarely available, studies indicate that all Sphyrnidae have declined by 89% in the Northwest and West Central Atlantic (Baum *et al.* 2003; Hayes *et al.* 2009) and by 99% in the Mediterranean (Ferretti *et al.* 2008). Abundance of the great hammerhead has declined by approximately 80% off the coast of South Africa (Dudley & Simpfendorfer 2006) and in the Eastern Atlantic (Camhi *et al.* 2009). Although little is known about great hammerhead stock structure, migration patterns, breeding or pupping grounds, mating systems, and other basic biologic information necessary to develop scientifically based conservation and management programs, several conservation measures have been undertaken in response to sustainability concerns. The great hammerhead has been assessed as globally endangered by the IUCN Redlist (Denham *et al.* 2012). In late 2010 the International Commission for the Conservation of Atlantic Tunas (ICCAT) prohibited the retention or sale of any body part or whole carcass of members of the family Sphyrnidae within the Atlantic Ocean basin and adjoining seas, and a CITES Appendix II listing for the great hammerhead will become effective September 2014.

Here I report a study of the population genetic structure of this endangered globally distributed cosmopolitan shark. I use sequences from the complete mitochondrial control region and genotypes from 12 microsatellite loci to evaluate population differentiation, contemporary and historic migration, and demographic patterns across a large portion of the species distribution. Comparison of markers with

different inheritance modes and speed of evolution allows for a more thorough examination of the overall divergence between populations and evaluation of the timing and extent of hybridization than studies utilizing a single marker. In this study I (1) evaluate the population genetic structure and infer demographic history of this endangered species, (2) investigate the existence of hybrid zones which might explain the genealogic discordance suggested by other studies, and (3) evaluate the possibility of utilizing standing genetic variation to identify the geographic source of market fins.

Materials and methods

Samples and DNA extraction

I analyzed 312 great hammerhead specimens from 9 collection areas in 3 oceanic basins, final sample numbers used in analyses are shown in Figure 1. Samples were obtained from several sources including other researchers, fisheries observer programs, and artisanal and recreational fishers. The collection areas and their geographical abbreviations used herein are: (i) Western North Atlantic (WNA): US Atlantic coast (USAtl), Gulf of Mexico (GoM), and Caribbean (Cbn); Western South Atlantic (WSA); Western Atlantic (WAtl): WNA and WSA combined; (ii) Northern Indian Ocean (NIO): Red Sea (RS) and Persian Gulf (PG); (iii) Australia (Aus): Indonesia (IND), western Australia (WAus) and eastern Australia (EAus). I analyzed an additional 31 samples obtained from the Hong Kong fin markets with geographic origins suggested by the trader records. Tissue samples were obtained as fin clips or muscle punches and preserved in 95% ethanol. Genomic DNA was isolated from tissue samples using the QIAGEN DNeasy Tissue kit (QIAGEN, Inc., Valencia, CA).

DNA sequencing

The entire mitochondrial control region (mtCR) plus some flanking DNA was amplified from a total of 277 sharks by primers CRF6 (5' AAG CGT CGA CCT TGT AAG TC 3') (Testerman, Chapter 1) and DasR2 (5’ GCT GAA ACT TGC ATG TGT AA 3’) (V. Richards, unpublished). PCR reactions were performed in 50 μl volumes containing 1μl of unquantified genomic DNA, 200 μM of each dNTP, 12.5 pM each primer, 1 U HotStar *Taq*^{TM} DNA polymerase (Qiagen, Hilden, Germany) and 5μl 10x *Taq*^{TM} reaction buffer.

PCR cycling conditions were 95ºC for 15 minutes, 35 cycles of 94ºC for 1 minute, 50ºC for 1 minute and 72ºC for 2 minutes, with a final extension of 5 minutes at 72ºC. In each set of PCR amplifications, a negative control with no genomic DNA was included to check for contamination. Amplified products were purified using the QIAquick PCR Purification Kit (QIAGEN, Inc.) prior to direct cycle sequencing with BigDye 3.1 Terminator chemistry (Applied Biosystems, Inc., Foster City, CA) of both strands using primers CRF6, DasR2, and two internal primers designed for this study, GhhCRF3 (5' CTC AGG TAG ACT TGA ACT ATC CTC G 3') and GhhCRR1 (5' GGA TAG TTC AAG TCT ACC TGA GTG TTC 3'). Sequencing reactions were purified using Dyex 2.0 Spin Kit (QIAGEN, Inc.) and sequenced on an AB3130 genetic analyzer (Applied Biosystems, Inc.). Sequences were aligned with GENEIOUS (Drummond *et al.* 2008) and alignments were checked and finalized by eye. Novel haplotype sequences are available from GenBank (accession numbers --- to ---).

Microsatellite genotyping

I genotyped 316 great hammerhead individuals and an additional 31 Hong Kong market derived samples from this species at 12 previously published nuclear microsatellite loci (Nance *et al.* 2009; Testerman *et al.* In preparation). Annealing temperatures (Ta) for each locus are provided in Supplementary Information Table S1. Microsatellite loci were amplified using forward primers with fluorescently labeled M13(-21) attached to their 5′ ends (Schuelke 2000). PCR reactions were performed in 12 µl volumes containing 1µl of unquantified genomic DNA, 200 µM of each dNTP, 2 pM forward primer, 5 pM reverse primer, 5 pM M13 tail, 1.5mM $MgCl_2$, 0.5 U HotStar *Taq*™ DNA polymerase (QIAGEN, Inc.) and 1.25µl 10x *Taq*™ reaction buffer. The few exceptions were 1.8 mM $MgCl_2$ for loci Ghh-A6-2, Ghh-C3-2 and Gh-D10-3, with locus Ghh0D10-3 requiring 5 pM forward primer and 10 pM each of the reverse primer and the M13 tail. PCR cycling conditions were 95ºC for 15 minutes, 35 cycles of 94ºC for 1 minute, the locus specific annealing temperature for 1 minute and 72ºC for 1 minute, with a final extension of 5 minutes at 72ºC. Fragments were separated on an AB 3130 genetic analyzer (Applied Biosystems, Inc.). Genotypes were scored using GENEMAPPER software version 3.7 (Applied Biosystems, Inc.) by comparison with the internal size standard LIZ 500

(Applied Biosystems, Inc.). Approximately 5% of the genotypes were re-amplified and re-scored to ensure genotyping repeatability and quality.

Genotypes were checked for null alleles, scoring errors, large allele dropout, and duplicates using MICROCHECKER (Van Oosterhout *et al.* 2004) and MICROSATELLITE TOOLKIT (Park 2001) for Microsoft EXCEL. Individual samples that failed to amplify at more than 3 loci were excluded from further analyses. Samples with identical multilocus genotypes or with up to two allele mismatches were considered to potentially be from the same individual; these genotypes were re-evaluated and corrected as appropriate. Several juveniles and individuals of unknown age were sampled at various collection areas. Because juvenile sharks may use the same nursery site for a few years (Wiley & Simpfendorfer 2007; Chapman *et al.* 2009), it is possible that these juveniles may have been either full or half siblings. I used the pairwise hypothesis testing option in ML-RELATE (Kalinowski *et al.* 2006) to identify related individuals (full sibs, half sibs, or parent-offspring) in each collection area. Likelihood ratios of 10 000 random dyads were simulated, and the hypothesis that a pair was related was accepted when the probability of their likelihood ratio was <0.01. When duplicate or related samples were found, one individual in each pair was randomly selected for inclusion in further analyses of mtDNA sequences and nuclear genotypes.

Statistical analysis

Due to small sample sizes from the western South Atlantic and Indonesia, data from those collection areas were not used in statistical or population-level analyses but were included in individual level analyses. For all statistical analyses with multiple tests, I evaluated significance levels after sequential Bonferroni correction (Rice 1989).

Analysis of mitochondrial DNA genetic diversity and population differentiation

Genetic diversity was evaluated using the ARLEQUIN software v3.5.1.2 (Excoffier & Lischer 2010) to calculate the number of haplotypes (*nh*), haplotype diversity (*h*) and nucleotide diversity (π) from the 7 collection areas with sample sizes of $n > 10$ samples. I investigated genetic structure within a geographic context by an analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) with the Tamura & Nei (Tamura & Nei 1993; Bowen *et al.* 2005) distance as implemented in ARLEQUIN. AMOVA calculates analogs

to Wright's F-statistics (Wright 1951, 1965), designated Φ_{ST}, based on the allelic content of haplotypes and haplotype frequencies. I initially calculated pairwise Φ_{ST} values among the seven collection areas. Based on the results obtained, samples from neighboring localities that had non-significant pairwise Φ_{ST} values ($P > 0.05$) were pooled for subsequent global population-level analyses. The amount of genetic differentiation within and among these pooled populations was evaluated through hierarchical AMOVA. A significance threshold for the covariance components was estimated using 1 000 nonparametric permutations.

I visualized patterns of differentiation among individuals and determined the presence of distinct genetic clusters by two different methods. First, principal coordinate analysis (PCoA) as implemented in GENEALEX v6.5b3 (Peakall & Smouse 2006; Peakall & Smouse 2012) was performed. Eigenvectors were calculated from a covariance matrix and the first two coordinates were plotted. Gene genealogies were estimated by constructing unrooted parsimony-based haplotype networks using the Templeton *et al.* (1992) method as implemented in TCS v1.21 (Clement *et al.* 2000) with the connection limit set initially at 95%. Ambiguities in the networks were resolved using criteria based on coalescent theory (Crandall & Templeton 1993; Pfenninger & Posada 2002); alternate connections are shown as dashed lines. To facilitate these analyses, haplotypes were identified using DNACOLLAPSER v 1.0 (Fredsted 2006).

Mitochondrial DNA population demographics analyses

A variety of approaches were used to evaluate population demographics and examine concordance between statistical and coalescent methods. As generation time and mutation rates are required to translate parameter estimates into demographic terms and because this information is lacking for the great hammerhead, I used the generation time and mtCR mutation rate (μ, the substitution rate per lineage per site per generation) estimated for its congener *Sphyrna lewini* by (Nance *et al.* 2011). First, summary statistics were calculated in ARLEQUIN, including Tajima's D statistic (Aris-Brosou & Excoffier 1996; Tajima 1996) which tests the hypothesis of selective neutrality and population equilibrium, and Fu's F_S (Fu 1997) which evaluates neutrality of mutations and is sensitive to population demographic expansion that leads to large negative values of F_S. For comparison purposes, the D* and F* statistics of Fu and Li (1993) were

calculated in DNASP v5.1 (Librado & Rozas 2009). Population expansion is suggested when Fu's F_S is significant and D* and F* statistics are not whereas the reverse combination suggests selection (Fu 1997). I also conducted a mismatch distribution analysis (Schneider & Excoffier 1999) in ARLEQUIN, and derived distribution goodness of fit parameters (specified below). The mismatch analysis computes the number of differences between pairs of haplotypes and uses a non-linear least-squares approach to estimate parameters of a sudden demographic (Rogers & Harpending 1992) or geographic (Ray *et al.* 2003) expansion. Goodness of fit was assessed by the sum of squares deviations (SSD) between the observed and expected mismatch distribution which tests the validity of the expansion model, and Harpending's raggedness index (*Hri*) (Harpending 1994) which tests the fit of the mismatch distribution. *Hri* provides indications of population expansion, with larger values of *Hri* typical of multimodal distributions found in stationary populations and smaller, non-significant values for unimodal or smoother distributions typical of expanding populations. Tau (τ) is calculated in the mismatch analysis and is a relative measure of the time in generations since population expansion ($\tau = 2\mu t$). I also report time since expansion in years (Time) calculated by $t = \tau/2\mu$ where μ is the mutation rate per site per generation. For the geographic expansion model, m is the migration rate between the sampled deme and a population of infinite size after T generations, $M = 2N_e m$ and N_e is the effective population size.

I estimated Θ ($\Theta = N_e\mu$) and growth rates for genetically differentiated populations using the Bayesian implementation in LAMARC v2.1.8 (Kuhner 2006), a coalescent-based MCMC method. I report N_e to facilitate comparison with N_e derived from nuclear microsatellite data, female effective population sizes can be calculated by $\Theta=2N_{ef}\mu$. I report directional migration rates between population pairs in terms of the biologically significant parameter $4N_em$ ($4N_em = M\Theta_{rec}$ where M is the LAMARC migration output parameter and Θ_{rec} is Θ of the receiving population) because the force of migration becomes strong enough to overcome genetic drift when 4Nem > 1 (Kuhner 2006). Initial data runs were conducted with wide priors to determine appropriate upper bounds, after which priors were revised for final data runs. Each data set was analyzed at least three times with different random seeds and run length of between 1.0×10^7 and

1.2 x 10^7 generations, while the final input prior values and all other inputs were held constant to evaluate convergence. For each of the output parameters, the means from each of the final runs were averaged to obtain the best estimate.

Finally, I inferred changes in N_e over time and estimated current N_e through Bayesian skyline plots (BSP) as implemented in BEAST v1.7.4 (Drummond *et al.* 2012). Three separate analyses were run for each genetically differentiated population using the HKY+I+G nucleotide substitution model, a constant skyline model with 10 groups, genealogies and parameters sampled every 1000 iterations, and automatic optimization of operators. The analyses were run for 1 x 10^8 iterations with a strict molecular clock set at the *S. lewini* derived mutation rate (Nance *et al.* 2011). Convergence was assumed when the ESS for each parameter in each replicate run exceeded 200. Skyline plots were generated in TRACER v1.5 (Drummond & Rambaut 2007).

Analysis of microsatellite genetic diversity and population differentiation

Input files were prepared for microsatellite data analyses using the software CREATE v. 1.36 (Coombs *et al.* 2008). Microsatellite loci were checked for deviations from Hardy-Weinberg equilibrium and evidence of linkage disequilibrium using GENEPOP v4.1 (Rousset 2008). Exact tests of Hardy-Weinberg equilibrium (HWE) were performed using the Markov chain method (Guo & Thompson 1992; Rousset & Raymond 1995), with 10 000 dememorizations, 20 batches and 5 000 iterations per batch. Global tests across loci and collection areas used Fisher's method. The software GENETIX v4.05.2 (Belkhir *et al.* 1996) was used to estimate observed (H_o) and expected (H_e) heterozygosity, number of alleles, and inbreeding coefficients (F_{IS}). I used FSTAT v2.9.3.2 (Goudet 1995, 2001) to calculate allelic richness (A_r) (El Mousadik & Petit 1996). The frequency of null alleles was estimated using the Expectation Maximization (EM) algorithm of (Dempster *et al.* 1977) as implemented in GENEPOP .

Pairwise F_{ST} values were calculated using GENEPOP to assess population-level genetic differentiation among all seven original collection areas with sample sizes >10 individuals. Fisher's exact tests for *P* values of genic differentiation were performed using the Markov chain method, with 10 000 dememorizations, 20 batches and 5 000 iterations per batch. Global tests across loci and collection areas used Fisher's method. Neighboring collection areas not significantly differentiated from each other were pooled

into putative populations for subsequent statistical and population-level analyses. Individual-based population differentiation was assessed using the following three approaches. First, principal coordinate analyses (PCoA) were performed in GENEALEX to visualize genetic differentiation among samples as described above. I then assessed genetic population structure using Bayesian clustering to assign individuals to populations in the program STRUCTURE version 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003), using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation. STRUCTURE analyses were performed using the collection areas as prior information, the admixture model, correlated allele frequencies, and a burn-in period of 100 000 MCMC generations followed by 200 000 iterations. The number of populations ranged from $K = 1$ through $K = 10$ with 10 replicates for each K. STRUCTURE HARVESTER v0.6.93 (Earl & vonHoldt 2012) was used to determine the most likely number of distinct genetic clusters by evaluating the logarithm of the probability of the data ($\ln P(D \mid K)$) (Pritchard *et al.* 2000) and estimates of ΔK (Evanno *et al.* 2005). Each individual's admixture proportions were averaged over the 10 replicates for the most likely K using the program CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007), and the output graphically displayed using the program DISTRUCT v1.1 (Rosenberg 2004). Finally, for comparative purposes, a third algorithm, FLOCK v2.0, (Duchesne & Turgeon 2009; Duchesne & Turgeon 2012) that uses an iterative reallocation, non-MCMC method was used to assign individuals into K distinct populations without any *a priori* sample location information. This program was run starting with $K = 2$ and increasing K until one stopping condition was reached, with 50 runs per K and 20 iterations per run. STRUCTURE may have more power to detect population differentiation when migration rates are low while FLOCK may have more power under conditions of high, sustained migration (Duchesne & Turgeon 2012).

Microsatellite population demographic analyses

Population demographics for the genetically differentiated populations identified by pairwise F_{ST} analyses of nuclear microsatellite data were assessed using a variety of approaches. Given the IUCN Endangered listing for great hammerheads, I tested for changes in population size using several methods differing in their underlying demographic models. The existence of bottlenecks in each of the genetically distinct

populations was assessed using two statistical moment-based estimators. The program BOTTLENECK v1.2.02 (Piry *et al.* 1999) calculates statistics that test for departures from equilibrium patterns of heterozygosity that can be disrupted when the effective population size changes significantly. The Wilcoxon sign-rank deficiency test in the BOTTLENECK program was used because it is the most powerful when fewer than 20 loci are analyzed (Piry *et al.* 1999). Calculations were performed assuming three different microsatellite mutation models: the infinite alleles model (IAM), the single-step model (SSM), and the two phase model (TPM). For the TPM, variance was set to 12 and 95% single step mutations were selected as recommended by the software authors. Significance was assessed over 10 000 replicates. I also attempted to detect population bottlenecks using the M-ratio test (Garza and Williamson 2001). The *M*-ratio statistic was calculated using the software *M_P_val* (Garza & Williamson 2001) and compared to a simulated equilibrium distribution calculated in the program *critical_M* selecting the TPM, 10 000 replicates, with the percentage of mutations that follow the single step model ($p_s = 0.9$) and the mean size of larger mutations ($\Delta_g = 3.5$) as suggested by the software authors. Critical values (M_C) were calculated for a range of θ (0.01, 0.1, 1 and 10), and statistical significance was assessed over 10 000 replicates.

Changes in great hammerhead population size were also assessed using coalescent methods. The hierarchical Bayesian MCMC model implemented in MSVAR 1.3 (Storz & Beaumont 2002) was used to estimate parameters including the current effective population size (N_0), the ancestral population size (N_1), the time of population size change (t), and the mutation rate (μ). This estimated microsatellite mutation rate was used in further demographic analyses. MSVAR program parameters used were the exponential growth model, 5 replicate runs with varied priors and broad hyerpriors for the model parameters, and run lengths of 2×10^9 or 3×10^9 steps with output reported every 100 000 or 150 000 steps. After removing a 10% burn-in, convergence was assessed by calculating the Gelman-Rubin multivariate scale reduction factor (PRSF) and the effective sample size (ESS) across the 5 independent runs using the CODA package in the software R v2.15.1 (R Development Core Team 2012). Estimates of the mode and 95% highest posterior density (HPD) intervals were obtained using the modeest v1.14 and coda v0.15-2 packages, respectively, in R.

Finally, the microsatellite data were analyzed with LAMARC and BEAST to allow comparison of population demographic history from both mtCR and microsatellite data using the same software. Population Θ ($\Theta=4N_e\mu$) and growth rates for genetically differentiated populations, and per-generation migration rates ($m=M*\Theta_{rec}$) between these populations were estimated using the Bayesian implementation in LAMARC. Initial data runs were conducted with wide priors to determine appropriate upper bounds and the priors were revised for final data runs. As suggested by the software authors, each data set was run for one extremely long final run of 3×10^7 generations. Changes in effective population size over time as well as a point estimate of the current effective population size were inferred through extended Bayesian skyline plots (EBSP) as implemented in BEAST. Three separate analyses were run for each population using the equal rate proportionality, unbiased mutational estimate, single phase, and a strict molecular clock with uniform priors (0, 20). The analyses were run for the maximum of 2.147×10^9 iterations. Convergence was assumed when the ESS for each parameter in each run exceeded 200.

Admixture analyses

Given the high mobility capabilities of large sharks and global distribution of great hammerheads, some level of recent gene flow between populations is possible. To assess gene flow, I estimated the number of first, second and third generation migrants between genetically differentiated populations as defined by microsatellite F_{ST} analyses using the Gensback = 2 options in STRUCTURE with the run parameters described above; using K = 2 based on the two populations identified by previous STRUCTURE analyses (western Atlantic and Northern Indian Ocean / Indo West Pacific or NIO/IWP). In these analyses, STRUCTURE assesses the proportion of ancestry (q) in both possible populations (WAtl or NIO/IWP). Individuals that are not migrants or admixed should have a posterior probability of being from the assumed population that is close to 1.0, with lower values indicating a migrant or admixed individual. I chose a cut-off value of 0.8 such that if q was < 0.8 for the sampling locality, that individual was assigned to the class with the highest posterior probability. Individuals were assigned as either a migrant or as having at least one migrant parent (F1) or grandparent (F2). The number of first generation migrants was analyzed using the Bayesian individual assignment method implemented in

GENECLASS 2.0 (Piry *et al.* 2004) at $p < 0.05$. Selected program parameters included the L=L_Home option as not all possible source populations were sampled, the Rannala and Mountain (1997) Bayesian criteria, and Monte-Carlo resampling (Paetkau *et al.* 2004) with 10 000 simulated individuals ($\alpha = 0.05$) to assess probabilities. Finally, because gene flow between the western Atlantic and the NIO/IWP populations could have resulted in individuals with hybrid nuclear genotypes, the probability of individuals belonging to distinct hybrid or purebred classes was estimated in the software NEWHYBRIDS v1.1beta (Anderson & Thompson 2002). The classes used were pure western Atlantic, pure NIO/IWP, F1 hybrid, F2 hybrid, backcross with western Atlantic and backcross with NIO/IWP. Because the northern Indian Ocean is statistically differentiated from Australia, and NEWHYBRIDS can only compare two parental populations, similar analyses were also run for the pairs of western Atlantic and northern Indian Ocean, western Atlantic and IWP, and northern Indian Ocean and IWP. NEWHYBRIDS analyses were run 10 times with default parameters, Jeffrey's priors, and a minimum burn-in of 20,000 steps with at least 200,000 additional steps. To be considered part of a class, an individual had to be assigned a posterior probability of at least 0.75.

Geographic origin of market fins

Both mtCR sequences and nuclear microsatellite genotypes were used to determine the broad geographic origin of 31 samples previously identified as originating from great hammerheads and obtained from the Hong Kong fin market (Abercrombie *et al.* 2005). To determine which population each market sample clustered with, the control region sequences and microsatellite genotypes were included in the PCoA analyses, the TCS haplotype network and the STRUCTURE runs as appropriate for the marker type. Additionally, GENECLASS was used to calculate the probability of an individual sample originating from one of the three statistically differentiated populations: the western North Atlantic, northern Indian Ocean or Australia.

Results

Genetic variation

Sequencing the complete mtCR (1 082 nucleotides) of 272 individuals identified 52 polymorphic sites that defined 90 haplotypes. Overall haplotype and nucleotide diversity were high at 0.8975 and 0.0108 respectively, with much higher diversity observed in the Indo West Pacific compared to the other collection areas. (Table 1a).

Multilocus genotypes were determined for a total of 264 individuals at 12 microsatellite loci after removing duplicate multilocus genotypes (n = 1), full siblings (n = 5) and those that amplified at less than 75% of loci (n = 15). Polymorphism ranged from 2 to 43 alleles per locus (mean = 10.08; Table 2), with no evidence of scoring errors or large allele dropout. A few loci exhibited lower than expected levels of heterozygosity in some collection areas, but there was no systematic pattern indicative of null alleles with only 4 loci exhibiting null allele frequencies > 0.10 (GhhC32 in US Atlantic, GhhD9 in US Atlantic and Gulf of Mexico, and Sle81 in the Persian Gulf). Similarly, some locus pairs showed evidence of linkage disequilibrium in a few collection areas, but none showed linkage disequilibrium consistently across collection areas. Although two loci had F_{IS} values that were significant in a single collection area (GhhC32 in the US Atlantic and GhhD9 in the Gulf of Mexico), none had significant F_{IS} values consistently across collection areas. Based on the above marker performance, all 12 loci were retained in the statistical analyses. Most loci had slightly higher levels of allelic richness and expected heterozygosity in Australia and the northern Indian Ocean than in the western Atlantic.

Genetic structure – Mitochondrial DNA

Pairwise mtCR Φ_{ST} values were not significantly different (data not shown) among collection areas within the western North Atlantic (USAtl, GoM, Cbn), and between western and eastern Australia (WAus, EAus), so those collection areas were pooled to comprise two statistically differentiated populations (i.e., WNA and IWP, respectively) in further population-level analyses. Pairwise Φ_{ST} values were significant among the western North Atlantic (pooled), Red Sea, Persian Gulf, and western/eastern Australia (pooled) collection areas (Table 3). AMOVA revealed that 84.71% of the variation exists

among these four populations with 15.29% of the variation within populations ($P < 0.00001$).

The mtCR haplotype statistical parsimony network revealed two divergent clades (Figure 2). One clade (WAtl/NIO) comprised haplotypes sampled in the US Atlantic, Gulf of Mexico, Caribbean Sea, western South Atlantic, Red Sea and Persian Gulf, including shared and very closely related haplotypes between the Western Atlantic Ocean and the Red Sea and Persian Gulf samples. The second clade (IWP) comprised haplotypes sampled from western and eastern Australia, as well as the single individual sampled in Indonesia. Great hammerheads from eastern and western Australia shared haplotypes; no haplotypes were shared between the two clades. The two clades, separated by 18 mutational steps, could only be joined at the 92%, rather than the traditional 95% confidence level used in TCS network construction. The WAtl/NIO clade contained a total of 28 haplotypes with a central, high frequency haplotype shared by 83 individuals that is connected to several single (1 individual) or low frequency haplotypes (shared by 2 – 4 individuals). The other four high frequency haplotypes (shared by 5 – 27 individuals) were nested around the central haplotype, and single mutational steps generally separated all haplotypes. In contrast, most of the 72 haplotypes in the IWP clade occurred in low frequency with only six high frequency haplotypes shared by 5 - 11 individuals.

Principal coordinate analysis of the mtCR sequences revealed two distinct genetic clusters, one comprising samples from the US Atlantic, Gulf of Mexico, Caribbean, Red Sea and Persian Gulf, the other cluster comprising western and eastern Australia samples (Figure 3a). PCoA Axis 1 separated individuals based on these geographic regions and explained 90.35% of the variation. Axis 2 explained 3.8% of the variation and further divided individuals into smaller clusters although there was no evidence of geographic structure within these smaller clusters.

Genetic structure – Microsatellite DNA

Microsatellite pairwise F_{ST} values were not significantly different among collection areas within the western North Atlantic, between the Red Sea and Persian Gulf, and between western and eastern Australia (not shown). However, pairwise F_{ST} values (Table 3) were significantly different among the pooled western North Atlantic

(WNA), pooled northern Indian Ocean (NIO) and pooled Australia (IWP) samples; the F_{ST} between the NIO and IWP samples is an order of magnitude smaller than between the WNA and NIO or WNA and IWP. Thus great hammerheads from these three regions were treated as statistically differentiated populations in further population-level analyses.

In the individual-based analyses, PCoA showed two slightly overlapping clusters (Figure 3b): one comprising samples from the US Atlantic, Gulf of Mexico, Caribbean, and western South Atlantic and a second cluster of individuals from the Red Sea, Persian Gulf, Indonesia and western and eastern Australia. PCoA Axis 1 explained 35.5% of the variation, Axis 2 explained 17.7%, and Axis 3 explained 14.1% of the variation. Bayesian STRUCTURE (Figure 4) and the FLOCK analyses also demonstrated the presence of the same k = 2 genetically distinct clusters revealed by PCoA. Hierarchical STRUCTURE and FLOCK analyses of only Red Sea, Persian Gulf, Australia and Indonesian samples returned k = 1 (not shown). Like the PCoA, STRUCTURE and FLOCK also grouped great hammerheads from the Red Sea and Persian Gulf with those from Australia / Indonesia.

Demographic analyses

Mitochondrial neutrality statistics (Table 4a) indicated demographic expansion and selective neutrality of the pooled WNA population, with Fu's F_S negative and significant, Tajima's D generally non-significant, and Fu and Li's D* and F* non significant. However, Tajima's D and Fu's F_S were inconsistent in their support of demographic expansion for the Red Sea and IWP populations. For all populations, the mismatch distribution fit parameters (SSD and *Hri*) did not differ significantly from that expected under a model of demographic expansion (Table 4b) or geographic expansion (Table 4c). Time since expansion (Time) was greater for the Australian population than the western North Atlantic, Red Sea and Persian Gulf populations. Estimates of θ by both BEAST and LAMARC (Table 5a) were highest for Australia, intermediate for the western North Atlantic and Red Sea, and lowest for the Persian Gulf, although absolute point estimates varied between the two programs. Population growth rates estimated in LAMARC were highest for the Red Sea, intermediate for the western North Atlantic and Persian Gulf, and lowest for Australia. Migration rate estimates from LAMARC were highest within the

western North Atlantic and northern Indian Ocean clusters and much lower between the Australian and western North Atlantic and northern Indian Ocean clusters. Directionality assessments indicated migration from the western North Atlantic to the northern Indian Ocean and from the Persian Gulf to the Red Sea. Bayesian skyline plots indicated slight population growth in the western North Atlantic, Red Sea and Persian Gulf, with rapid population expansion in Australia beginning in the Pleistocene (Figure 5).

Nuclear summary statistics generally showed mixed evidence of population bottlenecks (Table 5b). The Wilcoxon test in the program BOTTLENECK was only significant in all three populations under the IAM mutational model, the other mutation model results were non-significant. In contrast, the observed M value was smaller than the simulated critical value for all three populations and the pooled global sample, indicating that all have undergone bottlenecks. The only exception was that in all four cases, the observed value was larger than the critical value when $\Theta = 10$, indicating no population decline if the historic effective population size was large.

The MSVAR analyses (Table 5c) indicate large declines in each of the three populations, with point estimates for current effective population (N_0) size two to three magnitudes smaller than ancestral (N_1) effective size; however, there was considerable overlap in the HPDs surrounding these point estimates. The time since the population decline estimated by MSVAR indicates that these putative declines likely occurred in the Pleistocene in the northern Indian Ocean and Australia and during the Pliocene in the western North Atlantic. MCMC diagnostics indicate that the chains were well mixed (PRSF < 1.6) and effective sizes were large (ESS > 200 for each parameter in each population). The average estimated mutation rate across all three populations was 1.41×10^{-6}. The BEAST estimates of θ are largest in the western Atlantic and the northern Indian Ocean, and smaller in Australia. Although these analyses were run for the maximum number of generations allowed in BEAST, the ESS values were less than 200 so these results should be interpreted cautiously.

The LAMARC estimates of θ in the northern Indian Ocean and in Australia are similar to those obtained from BEAST, but the LAMARC estimate of θ for the western Atlantic is an order of magnitude smaller than the BEAST estimate. Population growth rates estimated in LAMARC based on the microsatellite data were small or negative, and

the confidence intervals in the western North Atlantic and Australia spanned 0 indicating a population decline (Kuhner 2006). The largest estimated migration rate was from the western North Atlantic into the northern Indian Ocean, with a strong indication of migration from Australia into the northern Indian Ocean and a small but biologically significant rate from the western North Atlantic into Australia.

GENECLASS identified four first generation migrants, all sampled in the northern Indian Ocean (Table 6). Two were classified as being from the western North Atlantic, and had admixed genotypes. Interestingly, NEWHYBRIDS classified both as being F2 individuals, while one was identified as being a second generation immigrant and the other was classified as both a first and a second generation immigrant with equal probability in the STRUCTURE analysis. The other two GENECLASS migrants were identified as being from Australia and were not recognized as admixed in the STRUCTURE or GENECLASS analyses. In the NEWHYBRIDS analyses of all samples combined most individuals were strongly assigned to either the western North Atlantic or to the northern Indian Ocean / Australia with a relatively few admixed individuals. In contrast, analyses in which the northern Indian Ocean and Australia were separated indicated that all northern Indian Ocean and Australian individuals were admixed, even in comparisons with the western Atlantic, while western Atlantic individuals were not admixed. This indicates that like STRUCTURE, NEWHYBRIDS cannot differentiate between the northern Indian Ocean and Australian samples, so only results from NEWHYBRIDS analyses of all individuals combined are reported here. The STRUCTURE analyses were more conservative than the NEWHYBRIDS analyses in that all four individuals identified as admixed by STRUCTURE were also identified as hybrids by NEWHYBRIDS, but five of the nine individuals identified as hybrids by NEWHYBRIDS were identified as not admixed by STRUCTURE. It is likely that six of the admixed individuals belong to the F2 generation of hybrids while the other three probably represent hybrids backcrossed with Australian individuals. The presence of both types of crosses indicates that hybrids are fertile and thus reproductive barriers are not absolute.

Geographic origins of market fins

Of the 31 market samples, 11 did not meet the minimum successful amplification percentages or were determined to be duplicates based on their microsatellite genotype as

described in *Microsatellite genotyping* above, and were excluded from these analyses. Market samples were assigned to a source population when the probability was at least 0.70. Eighteen of the included 20 market samples were assigned consistently by all methods used with both nuclear and mitochondrial markers, 13 to the western Atlantic and 5 to Australia (Table 7). One of the other 2 market samples had an Atlantic type mtCR haplotype, but the microsatellite genotype appeared to be of mixed ancestry. The other market sample was assigned to the western North Atlantic population by all methods except GENECLASS, which calculated similar probabilities for all three populations.

Discussion

Information about the population genetic structure and demographic history of endangered species can be critical in the design of conservation and management plans. This study assessed the population genetic structure of the globally distributed great hammerhead shark. I found strong differentiation between the western Atlantic and Australia with no sharing of mitochondrial haplotypes, identified genealogical discordance in individuals sampled in the northern Indian Ocean, and determined unambiguously the geographic origin of 18 of 20 samples obtained from the international shark fin trade. However, I did not have samples from the eastern Atlantic or the central or eastern Pacific and this study thus lacks data from these important regions. Additionally, a small sample size for the Persian Gulf may have affected some statistical analyses. Nevertheless, this is the first study to delineate the population genetic structure of this endangered species.

Genetic diversity and population demography

Summary statistics of genetic diversity for both mtCR and microsatellites were high in all collection areas, in genetically differentiated populations, and in the pooled samples overall. Haplotype and nucleotide diversity were larger in Australia than in the northern Indian Ocean or western North Atlantic, indicating greater genetic diversity in Australia. Additionally, the WNA/NIO clade has a star-like phylogeny is typically seen in

populations that have recently expanded, while the IWP clade has a more linear topography typical of deeper evolutionary relationships.

The microsatellite mutation rate estimated by Msvar (1.41×10^{-6}) seems reasonable given that reported microsatellite mutation rates for fish are in the range of 10^{-4} to 10^{-5} (Shimoda *et al.* 1999; Yue *et al.* 2007), the reported range for other taxa is 10^{-2} to 10^{-5} (Ellegren 2004) and mitochondrial and nuclear DNA appears to mutate approximately an order of magnitude more slowly in sharks than in other taxa (Martin *et al.* 1992; Martin 1999). This calculated rate is slower than the microsatellite mutation rates used by Karl *et al.* (2011) ($1x10^{-4}$) and Nance *et al.* (2011) ($1x10^{-5}$ to $2x10^{-4}$), but those were selected by the authors from other taxa and not based on their data.

Mitochondrial and nuclear statistical demographic analyses yielded conflicting results, with mitochondrial data providing evidence of demographic and geographic expansions and nuclear data indicating population bottlenecks. The moment-based estimators do not indicate population contraction at all (BOTTLENECK), or if the historic population size was large (M-RATIO). However, the nuclear statistical demographic methods used are sensitive to more recent events while coalescent methods are sensitive to more ancient events. These patterns are also evident in the contrast in population growth rate estimates from the LAMARC analyses of nuclear (decline) and mitochondrial (strong expansion) data. It is important to note that with the exception of N_0 from MSVAR, the estimates of effective population sizes are greater for nuclear data than for mitochondrial data. This could be partially explained by the observation that some of the demographic estimates have broad and overlapping confidence intervals or HPDs, thus too much weight should not be placed on the absolute value of these point estimates. Also, using a faster microsatellite mutation rate as in Karl *et al.* (2011) and Nance *et al.* (2011) would decrease effective population size estimates based on nuclear data. Additionally, the microsatellite models employed in BEAST had only recently been implemented, and the LAMARC point estimates of θ (largest in Australia) make more biological sense. Nonetheless, the genetic diversity indices and coalescence times suggest an Indo-west Pacific origin of the species with subsequent dispersal into the northern Indian Ocean and western Atlantic.

Population structure and genealogic discordance

Statistical population-level and individual-based methods for both mtCR and microsatellites were concordant in identifying strong differentiation between the western Atlantic and Australian samples and indicating the presence of at least two genetically distinct groups in both mitochondrial and nuclear data sets. However, mitochondrial and microsatellite data yielded contradictory results from the Red Sea and Persian Gulf samples. Mitochondrial data indicated a closer evolutionary relationship of the northern Indian Ocean to the western Atlantic samples, while microsatellite genotypes showed a stronger affinity of northern Indian Ocean to Australian samples. The individual-based methods clearly show these relationships in the haplotype network, PCoA analyses and STRUCTURE plots, although the PCoA microsatellite plot shows some overlap of the western Atlantic and northern Indian Ocean / Australian individuals.

Statistical analyses revealed more detailed subdivisions among differentiated populations. Pairwise Φ_{ST} values from mtCR were significant among the western North Atlantic, Red Sea, Persian Gulf and Australia indicating four genetically distinct populations, and pairwise F_{ST} values from nuclear microsatellites indicated three genetically distinct populations. However, the pairwise Φ_{ST} values between Australia and each of the western North Atlantic, Red Sea and Persian Gulf were approximately four times larger than between the Persian Gulf and both the western North Atlantic and Red Sea, which are in turn an order of magnitude greater than between the western North Atlantic and Red Sea. Similarly, the pairwise F_{ST} values between the western North Atlantic and both the northern Indian Ocean and Atlantic are almost an order of magnitude greater than the northern Indian Ocean and Australia.

Low but statistically significant levels of genetic differentiation have been shown to be biologically meaningful and persistent over time in other marine fish (Knutsen *et al.* 2011), raising the issue of why the individual based and statistical methods result in different estimates of the number and location of genetically distinct populations. The difference between the statistical and individual based methods could be due to the inherent sensitivities of these methods, with the statistical methods teasing out more contemporary influences on population structure. This is supported by estimates of

coalescent times, which are generally more ancient for Australia and more recent for the northern Indian Ocean and western Atlantic.

What is clear across all analyses is that individuals from the Red Sea and Persian Gulf have western Atlantic type mitochondrial haplotypes but Australian-type microsatellite genotypes. The haplotype network supports ongoing long-term gene flow between the western Atlantic and northern Indian Ocean since there are multiple western Atlantic type haplotypes in northern Indian Ocean individuals as well as a few closely related novel haplotypes found only in Persian Gulf individuals. Migration rate estimates yield additional insight into this genealogical discordance. Migration rates greater than one individual per generation are biologically significant because they are sufficient to overcome the effects of genetic drift and prevent substantial genetic divergence (Wright 1931). Migration rates between the western Atlantic and northern Indian Ocean and between the Red Sea and Persian Gulf the Northern Indian Ocean are larger than those between the Northern Indian Ocean and Australia. Migration rates between the northern Indian Ocean and Atlantic are only biologically significant with nuclear microsatellites, while estimated rates between the western Atlantic and Australia are not significant from either mitochondrial or nuclear loci. Additionally, the directionality of these estimates indicates higher rates of gene flow into the northern Indian Ocean from both the western Atlantic and Australia than out of the northern Indian Ocean. The higher migration rates and more recent coalescent times between the western Atlantic and northern Indian Ocean indicate more connectivity between the western Atlantic and northern Indian Ocean, potentially facilitating homogenization of the mitochondrial haplotypes.

The northern Indian Ocean is a unique biogeographic location, due in part to the seasonal reversals of the Monsoon and Somali Currents resulting in alternating warm water and cold nutrient-rich waters. In a survey of coral reef fauna, Briggs and Bowen (2012) found that there are biogeographic breaks between the western Indian Ocean, the Red Sea Province, and the Polynesian Province (extending from the eastern Arabian Gulf to Polynesia and referred to herein as the IWP). The western Indian Ocean and the Red Sea Province both have high levels of endemism justifying their designation as two distinct biogeographic provinces separate from the IWP. Additionally, DiBattista *et al.* (2013) reported that the Red Sea has higher numbers of endemic reef fish than previously

thought and that at least some of these species have dispersed out of the Red Sea into the Indian Ocean. Perhaps it is not so surprising then to find a unique genetic pattern in this area. Indeed, a sister-species, *Sphyrna lewini*, has also been shown to have western Atlantic-type mtCR haplotypes and IWP-type microsatellite genotypes in some individuals sampled in the western Indian Ocean (Duncan *et al.* 2006; Daly-Engel *et al.* 2012).

This intriguing pattern of mito-nuclear discordance in two globally distributed sphyrnid species can be explained by dispersal patterns. Duncan *et al.* (2006) suggested an IWP origin for *S. lewini* followed by a westerly dispersal into the Indian and Atlantic Oceans, with a subsequent eastward migration from the Atlantic into the Indian Ocean. My migration rate and expansion time estimates suggest a similar pattern for *S. mokarran.* Peeters *et al.* (2004) demonstrated that a connection between the Atlantic and Indian Oceans through the Agulhas Current occurs consistently during inter-glacial periods. Not surprisingly, many highly vagile pelagic species including billfish (Alvarado Bremer *et al.* 2005), tuna (Chow & Ushiama 1995; Durand *et al.* 2005), sea turtles (Bowen & Karl 2007), marine mammals (Oremus *et al.* 2009; Amaral *et al.* 2012), and great white sharks (Gubili *et al.* 2010) exhibit a pattern of dispersal from the Indian to the Atlantic Ocean. Agulhas transport has also been documented for reef-associated species (Bowen *et al.* 2001; Lessios *et al.* 2001; Rocha *et al.* 2005). Most of these species have a larval stage and individuals are presumably transported passively through the Agulhas current as larvae. Eastward dispersal from the Atlantic into the Indian Ocean has been suggested for *S. lewini* (Duncan *et al.* 2006) and the green turtle *Chelonia mydas* (Bourjea *et al.* 2007), both of which lack a larval stage and are active swimmers from the time they enter the open ocean. This is the third report of such an anti-Agulhas dispersal, this time by a large strong swimming shark.

But if the Agulhas current is not a barrier to this large shark, then what could be causing the mito-nuclear discordance between the northern Indian Ocean and Australia? Many other sharks have been reported to show evidence of male-mediated dispersal, including two sphyrnid sharks, *S. lewini* (Daly-Engel *et al.* 2012) and *Sphyrna zygaena* (Testerman, Chapter 4). I suggest male mediated dispersal by the great hammerhead based on the following: the low but biologically significant migration rates between

Australia and the western Atlantic and northern Indian Ocean estimated from nuclear data combined with insignificant migration rates estimated from mitochondrial data, and by admixed individuals that have the expected oceanic haplotype but an admixed nuclear genotype. Petit and Excoffier (2009) reported that markers associated with the least dispersing sex are more introgressed, such that in species with male-mediated dispersal nuclear markers will be less introgressed than mitochondrial markers. During glacial maxima, lowered sea levels resulted in the formation of a land bridge between Asia and Australia (Voris 2000) that in conjunction with increased upwelling of cold waters may have formed a relatively impenetrable barrier (Flemminger 1986) to tropical and warm temperate species such as the great hammerhead. Many studies have shown a lasting genetic signature from decreased connectivity between the Indian and Pacific Oceans (Barber *et al.* 2006; Gaither *et al.* 2010; Gaither *et al.* 2011; Mirams *et al.* 2011; Phillips *et al.* 2011; Baums *et al.* 2012) due to this intermittent Indo-Pacific Barrier (IPB), particularly in coral reef associated species that may disperse more slowly than a large actively-swimming shark. I therefore raise the possibility that the hybrid zone in the NIO is an artifact of past climate change. Under this scenario, during glacial maxima great hammerheads in the Indian Ocean would be effectively separated from those in the Pacific Ocean, allowing time for genetic differentiation to occur. As sea levels rose during warmer interglacials great hammerheads could once again disperse along the entire Australian coastline. My results provide no evidence of female philopatry along continuous coastlines, so I would expect a relatively rapid reduction in differentiation in both mitochondrial and nuclear makers between western and eastern Australia as the IPB receded. However, my results also suggest that male-mediated dispersal would be more likely between Australia and the northern Indian Ocean giving rise to the observed mito-nuclear discordance in the northern Indian Ocean. Thus, a combination of male-mediated dispersal between Australia and northern Indian Ocean and past climatic change lends further support to my hypothesis of eastward dispersal via the Cape of Good Hope despite the cold waters of the prevailing Benguela Current.

Finally, I found no sharing of mtCR haplotypes or microsatellite genotypes between the western Atlantic and Australia. Naylor *et al.* (2012) suggested that similar divergence levels between the western Atlantic and IWP at mtNADH2 supports

recognition of these two populations as separate allopatric species. Indeed, my mtCR haplotype network only joined the western Atlantic and Australian lineages at the 92% confidence level, which may suggest the two populations are separate species. However, the mean rate of concordance between the number of taxa and the number of 95% sub-networks was considerably higher in studies that utilized COI than in all other studies (Hart & Sunday 2007). In fact, Wong *et al.* (2009) sequenced the mtCOI from a subset of my samples and found no geographic differentiation between the western Atlantic and Australia, with 28 of the 30 samples sharing a single haplotype that included western Atlantic and Australian individuals. I believe that the existence of reproductively successful hybrids, the relatively short time since divergence, and the lack of geographic differentiation at COI indicates that the western Atlantic, northern Indian Ocean and Australia represent genetically distinct populations of a single species. Further study with additional individuals from currently unsampled regions is necessary to fully evaluate the evolutionary history and taxonomic status of this species.

Geographic assignment of market samples

An important goal in fisheries management efforts is to manage species at the population or stock level (Dizon *et al.* 1992), however little to no traditional monitoring occurs at most shark landing sites resulting in a lack of species level catch rates. Previous genetic studies indicated that other mitochondrial sequences could be used to identify *S. mokarran* at the species level (COI, Ward *et al.* 2008; Wong *et al.* 2009) and distinguish between the Atlantic and Pacific populations (NADH2, Naylor *et al.* 2012). My results show that the observed patterns of genetic divergence among the western North Atlantic, northern Indian Ocean and Australia at the mtCR in *S. mokarran* are sufficient to allow identification of the broad geographic origin of market samples to one of the three major oceanic basins. Additionally, I found that 75% of the market samples included in this study originated in the western Atlantic, 25% originated in Australia, and none from the northern Indian Ocean. Thus genetic data can be utilized in conservation and management efforts, for example in estimating landings by population of origin, enabling enforcement of regional harvest quotas and bans, and monitoring the global catch and trade of this endangered species.

Conclusions

Based on mitochondrial control region sequence and microsatellite genotypes, there is strong genetic differentiation between the western Atlantic and Australian populations and a hybrid zone in the northern Indian Ocean. The genetic differentiation within the species is sufficient to determine the broad geographic origin of *S. mokarran* in trade. I uncovered hints of species origin and history, suggestive of an origin in the IWP with subsequent dispersal into the northern Indian Ocean and the Atlantic Ocean and recent migration from the western Atlantic back into the Indian Ocean. However, I lack samples from the eastern Atlantic, western Indian, and central and eastern Pacific Oceans and thus cannot infer the full evolutionary history of the species. Finally, the high levels of genetic diversity within and among populations imply that although large declines in abundance have been suggested there has not been a corresponding decrease in genetic diversity. This bodes well for the future recovery of this endangered species if conservation measures can be implemented quickly.

Acknowledgements

For help in sample collection I thank Debra Abercrombie, Jim Abernathy, Dareen Almojil, Edd Brooks, Jose Castro, Demian Chapman, Shelley Clarke, Andy Danylchuck, Sean Fitzpatrick, Mark Grace, Dean Grubbs, Simon Gulak, Lori Hale, Neil Hammerschlag, Ed Keith, James Lea, Tim Leary, Rick Martin, Cesar Martins, Rory McCauley, Phil Motta, Holly Nance, Todd Neahr, Danillo Pinhal, Paulo Prodhöl, Dennis Reid, M. Reinfandt, Eric Sanders, Colin Simpfendorfer, Buck Snelson, Ted Testerman, John Tyminski, Brad Wetherbee, Rachel Wilborn, Tanya Wiley, and Crispen Wilson. Financial support for this project was received from the Save Our Seas Foundation, Florida Sea Grant, Hai Siftung Shark Foundation and the Guy Harvey Ocean Foundation.

References

Abercrombie DL, Clarke SC, Shivji MS (2005) Global-scale genetic identification of hammerhead sharks: Application to assessment of the international fin trade and law enforcement. Conservation Genetics.

Alvarado Bremer JR, Vinas J, Mejuto J, Ely B, Pla C (2005) Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. Molecular Phylogenetics and Evolution, 36, 169-187.

Amaral AR, Beheregaray LB, Bilgmann K, Freitas L, Robertson KM, Sequeira M, Stockin KA, Coelho MM, Möller LM (2012) Influences of past climatic changes on historical population structure and demography of a cosmopolitan marine predator, the common dolphin (genus *Delphinus*). Molecular Ecology, 21, 4854-4871.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics, 160, 1217-1229.

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Moleclar Biology and Evolution, 13, 494-504.

Au DW, Smith SE, Show C (2009) Shark productivity and reproductive protection, and a comparison with teleosts. In: Sharks of the open ocean: biology, fisheries and conservation (eds. MD Camhi, EK Pikitch, EA Babcock). Blackwell Publishing Ltd., New Jersey.

Avise JC, Shapira JF, Daniel SW, Aquadro CF, Lansman RA (1983) Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. Moleclar Biology and Evolution, 1, 38-56.

Barber PH, Erdmann MV, Palumbi SR (2006) Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the coral triangle. Evolution, 60, 1825-1839.

Bastos-Silveira C, Santos SM, Monarca R, Mathias MdL, Heckel G (2012) Deep mitochondrial introgression and hybridization among ecologically divergent vole species. Molecular Ecology, 21, 5309-5323.

Baum JK, Blanchard W (2010) Inferring shark population trends from generalized linear mixed models of pelagic longline catch and effort data. Fisheries Research, 102, 229-239.

Baum JK, Kehler DG, Myers RA (2005) Robust estimates of decline for pelagic shark populations in the northwest Atlantic and Gulf of Mexico. Fisheries, 30, 27-29.

Baum JK, Myers RA (2004) Shifting baselines and the decline of pelagic sharks in the Gulf of Mexico. Ecology Letters, 7, 135-145.

Baum JK, Myers RA, Kehler DG, Worm B, Harley SJ, Doherty PA (2003) Collapse and conservation of shark populations in northwest Atlantic. Science, 299, 389-392.

Baums IB, Boulay JN, Polato NR, Hellberg ME (2012) No gene flow across the Eastern Pacific Barrier in the reef-building coral *Porites lobata*. Molecular Ecology, 21, 5418-5433.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire génome, populations, interactions, CNRS UMR, 5000, 1996-2004.

Benavides M*, et al.* (2011) Global phylogeography of the dusky shark *Carcharhinus obscurus*: implications for fisheries management and monitoring the shark fin trade. Endangered Species Research, 14, 13-22.

Borrero-Perez GH, Gonzalez-Wantuemert M, Marcos C, Perez-Ruzafa A (2011) Phylogeography of the Atlanto-Mediterranean sea cucumber *Holothuria mammata*: the combined effects of historical processes and current oceanographical pattern. Molecular Ecology, 20, 1964-1975.

Bourjea J, Lapegue S, Gagnevin L, Broderick D, Mortimer JA, Ciccione S, Roos D, Taquet C, Grizel H (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. 16, 175-186.

Bowen BW, Bass A, Rocha LA, Grant W, Robertson DR (2001) Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. Evolution, 55, 1029-1039.

Bowen BW, Bass AL, Soares L, Toonen RJ (2005) Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). Molecular Ecology.

Bowen BW, Karl SA (2007) Population genetics and phylogeography of sea turtles. Molecular Ecology, 16, 4886-4907.

Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. Journal of Biogeography, 39, 12-30.

Camhi MD, Valenti S, Fordham S, Fowler S, Gibson C (2009) The conservation status of pelagic sharks and rays: Report of the IUCN shark specialist group pelagic shark red list workshop.

Chapman DD, Babcock EA, Gruber SH, DiBattista JD, Franks BR, Kessel SA, Guttridge T, Pikitch EK, Feldheim KA (2009) Long-term natal site-fidelity by immature lemon sharks (*Negaprion brevirostris*) at a subtropical island. Molecular Ecology, 18, 3500-3507.

Chow S, Ushiama H (1995) Global population structure of albacore (*Thunnus alalunga*) inferred by RFLP analysis of the mitochondrial ATPase gene. Marine Biology, 123, 39-45.

Clarke SC, McAllister MK, Milner-Gulland EJ, Kirkwood GP, Michielsens CGJ, Agnew DJ, Pikitch EK, Hideki N, Shivji MS (2006) Global estimates of shark catches using trade records from commercial markets. Ecology Letters, 9, 1115-1126.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9, 1657-1659.

Compagno LJV (1984) FAO Species Catalog. Vol. 4 Sharks of the world, Part 2 - Carcharhiniformes. In: FAO Species Catalog. Food and Agriculture Organization of the United Nations, Rome.

Coombs JA, Letcher BH, Nislow KH (2008) CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. Molecular Ecology Resources, 8, 578-580.

Cortes E (2002) Incorporating Uncertainty into Demographic Modeling: Application to Shark Populations and Their Conservation. Conservation Biology, 16, 1048-1062.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics, 134, 959-969.

Cunha RL, Coscia I, Madeira C, Mariani S, Stefanni S, Castilho R (2012) Ancient Divergence in the Trans-Oceanic Deep-Sea Shark *Centroscymnus crepidater*. PLOS One, 7, e49196.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW (2012) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). PLOS One, 7, e29986.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution, 24, 332-340.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 1-38.

Denham J*, et al.* (2012) *Sphyrna mokarran*. In: 2012 IUCN Red List of Threatened Species. IUCN Red List. Downloaded on 13 November 2012.

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, Choat JH, Craig MT, Skillings DJ, Bowen BW (2013) After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. Journal of Biogeography, 40, 1170-1181.

Dizon AE, Lockyer C, Perrin WF, DeMaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. Conservation Biology, 6, 24-36.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7, 214.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2008) Geneious v4.0.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969-1973.

Duchesne P, Turgeon J (2009) FLOCK: a method for quick mapping of admixture without source samples. Molecular Ecology Resources, 9, 1333-1344.

Duchesne P, Turgeon J (2012) FLOCK Provides Reliable Solutions to the "Number of Populations" Problem. Journal of heredity, 103, 734-743.

Dudley SJ, Simpfendorfer CA (2006) Population status of 14 shark species caught in the protective gillnets off KwaZulu-Natal beaches, South Africa, 1978-2003. Marine and Freshwater Research, 57, 225-240.

Dulvy ND, Forrest RE (2010) Life histories, population dynamics, and extinction risks in chondrichthyans. In: Sharks and their relatives II: Biodiversity, adaptive physiology and conservation (eds. JC Carrier, JA Musick, MR Heithaus), pp. 635-676. CRC Press, Boca Raton, FL.

Duncan K, Martin A, Bowen B, DeCouet H (2006) Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*). Molecular Ecology, 15, 2239-2251.

Durand JD, Collet A, Chow S, Guinand B, Borsa P (2005) Nuclear and mitochondrial DNA markers indicate unidirectional gene flow of Indo-Pacific to Atlantic bigeye tuna (*Thunnus obesus*) populations, and their admixture off southern Africa. Marine Biology, 147, 313-322.

Dutton P*, et al.* (2013) Population stock structure of leatherback turtles (*Dermochelys coriacea*) in the Atlantic revealed using mtDNA and microsatellite markers. Conservation Genetics, 14, 625-636.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359-361.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. Theoretical applications in genetics, 92, 832-839.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics, 5, 435-445.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14, 2611-2620.

Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: A new series of programs to perform population genetic analyses under Linux and Windows. Molecular Ecology Resources, 10, 564-567.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131, 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics, 164, 1567-1587.

Ferretti F, Myers RA, Serena F, Lotze HK (2008) Loss of Large Predatory Sharks from the Mediterranean Sea. Conservation Biology, 22, 952-964.

Ferretti F, Worm B, Britten GL, Heithaus MR, Lotze HK (2010) Patterns and ecosystem consequences of shark declines in the ocean. Ecology Letters, 13, 1055-1071.

Flemminger A (1986) The Pleistocene equatorial barrier between the Indian and Pacific oceans and a likely cause for Wallace's line. In: Pelagic biogeography. UNESCO technical papers in marine science., pp. 84-97.

Fredsted PV (2006) FaBox - an online fasta sequence toolbox, http://www.birc.au.dk/fabox.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147, 919-925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics, 133, 693-709.

Gaither MR, Bowen BW, Bordenave TR, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT (2011) Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. BMC Evolutionary Biology, 11, 189.

Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). Journal of Biogeography, 37, 133-147.

Garcia VB, Lucifora LO, Myers RA (2008) The importance of habitat and life history to extinction risk in sharks, skates, rays and chimaeras. Proceedings of the Royal Society B, 275, 83-89.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology, 10, 305-318.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of heredity, 86, 485-486.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Gubili C, Bilgin R, Kalkan E, Karhan SU, Jones CS, Sims DW, Kabasakal H, Martin AP, Noble LR (2010) Antipodean white sharks on a Mediterranean walkabout? Historical dispersal leads to genetic discontinuity and an endangered anomalous population. Proc. R. Soc. B, online.

Guo SW, Thompson EA (1992) Performing exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics, 48, 361-372.

Harpending R (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology, 66, 591-600.

Hart MW, Sunday J (2007) Things fall apart: biological species form unconnected parsimony networks. Biology Letters, doi:10.1098/rsbl.2007.0307.

Hayes CG, Jiao Y, Cortés E (2009) Stock Assessment of Scalloped Hammerheads in the Western North Atlantic Ocean and Gulf of Mexico. North American Journal of Fisheries Management, 29, 1406-1417.

Heithaus MR, Frid A, Wirsing AJ, Worm B (2008) Predicting ecological consequences of marine top predator declines. Trends in Ecology & Evolution, 23, 202-210.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, 23, 1801-1806.

Jennings S, Kaiser MJ, Reynolds JD (2001) Marine fisheries ecology Blackwell Science.

Kalinowski ST, Wagner AP, Taper ML (2006) ml-relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes, 6, 576-579.

Karl SA, Castro ALF, Lopez JA, Charvet P, Burgess GH (2011) Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. Conservation Genetics, 12, 371-382.

Knutsen H, Olsen EM, Jorde PE, Espeland SH, Andre C, Stenseth NC (2011) Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. Molecular Ecology, 20, 768-783.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics, 22, 768-770.

Last PR, Stevens JD (2009) Sharks and rays of Australia, 2nd edn. Harvard University Press, Cambridge, Massachusetts.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. Evolution, 55, 955-975.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451-1452.

Martin A, Naylor G, Palumbi S (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature, 357, 1992.

Martin AP (1999) Substitution rates of organelle and nuclear genes in sharks: implicating metabolic rate (again). Molecular Biology and Evolution, 16, 996-1002.

Mirams A, Treml E, Shields J, Liggins L, Riginos C (2011) Vicariance and dispersal across an intermittent barrier: population genetic structure of marine animals across the Torres Strait land bridge. Coral Reefs, 30, 937-949.

Morgan A, Burgess GH (2007) At-vessel fishing mortality for six species of sharks caught in the Northwest Atlantic and Gulf of Mexico. Gulf and Caribbean Research, 19, 123-129.

Myers RA, Baum JK, Shepherd TD, Powers SP, Peterson CH (2007) Cascading effects of the loss of apex predatory sharks from a coastal ocean. Science, 315, 1846-1850.

Nance HA, Daly-Engel TS, Marko PB (2009) New microsatellite loci for the endangered scalloped hammerhead shark, *Sphyrna lewini*. Molecular Ecology Resources, doi: 10.1111/j.1755-0998.2008.02510.x.

Nance HA, Klimley P, Galvan-Magana F, Martinez-Ortiz J, Marko PB (2011) Demographic Processes Underlying Subtle Patterns of Population Structure in the Scalloped Hammerhead Shark, *Sphyrna lewini*. PLOS One, 6, e21459.

Naylor GJP, Caira JN, Jensen K, Rosana KAM, White WT, Last PR (2012) A DNA Sequence–Based Approach To the Identification of Shark and Ray Species and Its

Implications for Global Elasmobranch Diversity and Parasitology. Bulletin of the American Museum of Natural History, 1-262.

Oremus M, Gales R, Dalebout ML, Funahashi N, Endo T, Kage T, Steel D, Baker SC (2009) Worldwide mitochondrial DNA diversity and phylogeography of pilot whales (*Globicephala* spp.). Biological Journal of the Linnean Society, 98, 729-744.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology, 13, 55-65.

Palumbi SR, Cipriano F, Hare MP (2001) Predicting nuclear gene coalescence from mitochondrial data: The three-times rule. Evolution, 55, 859-868.

Park S (2001) Trypanotolerance in West African cattle and the population genetic effects of selection, PhD Dissertation, University of Dublin, Dublin, Ireland.

Patarnello T, Volckaert FAMJ, Castilho R (2007) Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break? Molecular Ecology, 2007, 4426-4444.

Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics, 28, 2537-2539.

Peakall ROD, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288-295.

Peeters FJC, Acheson R, Brummer G-JA, de Ruijter WPM, Schneider RR, Ganssen GM, Ufkes E, Kroon D (2004) Vigourous exchange between the Indian and Atlantic oceans at the end of the past five glacial periods. Nature, 430, 661-665.

Perez-Losada M, Nolte MJ, Crandall KA, Shaw PW (2007) Testing hypotheses of population structuring in the Northeast Atlantic Ocean and Mediterranean Sea using the common cuttlefish *Sepia officinalis*. Molecular Ecology, 16, 2667-2679.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. Trends in Ecology & Evolution, 24, 386-393.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration and secondary contact. Evolution, 56, 1776-1788.

Phillips NM, Chaplin JA, Morgan DL, Peverell SC (2011) Population genetic structure and genetic diversity of three critically endangered *Pristis* sawfishes in Australian waters. Marine Biology, 158, 903-915.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of heredity, 95, 536-539.

Piry S, Luikart G, Cornuet J-M (1999) BOTILENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of heredity, 90, 502-503.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA, 94, 9197-9201.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol., 20, 76-86.

Rice WR (1989) Analyzing tables of statistical tests. Evolution, 43, 223-225.

Richards VP*, et al.* (2013) Patterns of Population Structure for Inshore Bottlenose Dolphins along the Eastern United States. Journal of heredity, 104, 765-778.

Rocha LA, Robertson D, Rocha CR, Tassell JL, Craig MT, Bowen BW (2005) Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. Molecular Ecology, 14, 3921-3928.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution, 9, 552-569.

Rose DA (1996) An overview of world trade in sharks and other cartilaginous fishes, pp. 1-106. TRAFFIC International, Cambridge.

Rosenberg NA (2004) distruct: a program for the graphical display of population structure. Molecular Ecology Notes, 4, 137-138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8, 103-106.

Rousset F, Raymond M (1995) Testing heterozygote excess and deficiency. Genetics, 140, 1413-1419.

Roy D, Paterson G, Hamilton PB, Heath DD, Haffner GD (2007) Resource-based adaptive divergence in the freshwater fish *Telmantherina* from Lake Matano, Indonesia, pp. 35-48.

Ruegg K, Rosenbaum H, Anderson E, Engel M, Rothschild A, Baker CS, Palumbi S (2013) Long-term population size of the North Atlantic humpback whale within the context of worldwide population structure. Conservation Genetics, 14, 103-114.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. Genetics, 152, 1079-1089.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18, 233-234.

Shepherd TD, Myers RA (2005) Direct and indirect fishery effects on small coastal elasmobranchs in the northern Gulf of Mexico. Ecology Letters, 8, 1095-1104.

Shimoda N*, et al.* (1999) Zebrafish Genetic Map with 2000 Microsatellite Markers. Genomics, 58, 219-232.

Singhal S, Moritz C (2012) Testing hypotheses for genealogical discordance in a rainforest lizard. Molecular Ecology, 21, 5059-5072.

Storz JF, Beaumont MA (2002) Testing for genetic evidence of population expansion and contraction: An empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. Evolution, 56, 154-166.

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics, 143, 1457-1465.

Takahashi T, Hori M (2012) Genetic and Morphological Evidence Implies Existence of Two Sympatric Species in *Cyathopharynx furcifer* (Teleostei: Cichlidae) from Lake Tanganyika. International Journal of Evolutionary Biology, 2012.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Moleclar Biology and Evolution, 10, 512-526.

Templeton A, Crandall K, Sing C (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics, 132, 619-633.

Testerman CB, Fitzpatrick S, Prodohl PA, Shivji M (In preparation) Development and characterization of novel microsatellite loci for the great hammerhead shark *Sphyrna mokarran* and their cross-species amplification among other sphyrnid species.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535-538.

Voris HK (2000) Maps of Pleistocene sea levels in SoutheastAsia: shorelines, river systems and time duration. Journal of Biogeography, 27, 1153-1167.

Wares JP (2002) Community genetics in the Northwestern Atlantic intertidal. Molecular Ecology, 11, 1131-1144.

Warren BH, Bermingham E, Bourgeois Y, Estep LK, Prys-Jones RP, Strasberg D, Thébaud C (2012) Hybridization and barriers to gene flow in an island bird radiation. Evolution, 66, 1490-1505.

Welch AJ, Yoshida AA, Fleischer RC (2011) Mitochondrial and nuclear DNA sequences reveal recent divergence in morphologically indistinguishable petrels. Molecular Ecology, 20, 1364-1377.

Wiley TR, Simpfendorfer CA (2007) The ecology of elasmobranchs occurring in the Everglades National Park, Florida: implications for conservation and management. Bulletin of Marine Science, 80, 171-189.

Wong EH-K, Shivji MS, Hanner RH (2009) Identifying sharks with DNA barcodes: assessing the utility of a nucleotide diagnostic approach. Molecular Ecology Resources, 9, 243-256.

Wright S (1931) Evolution in mendelian populations. Genetics, 16, 97-159.

Wright S (1951) The genetical structure of populations. Annals of Eugenics, 15, 323-354.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, 19, 395-420.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). Genetica, 129, 329-331.

Tables

Table 1. Population statistics for *Sphyrna mokarran*. Number of individuals (*n*), number of haplotypes (*nh*), haplotype diversity (*h*), standard deviation (*SD*), nucleotide diversity (π). Collection areas are coded as follows: (i) Western North Atlantic (WNA): US Atlantic coast (USAtl), Gulf of Mexico (GoM), and Caribbean (Cbn); (ii) Northern Indian Ocean (NIO): Red Sea (RS) and Persian Gulf (PG); (iii) Indo West Pacific (IWP): western Australia (WAus) and eastern Australia (EAus).

	Western North Atlantic (WNA)			Pooled	North Indian Ocean (NIO)		Indo West Pacific (IWP)		Pooled	Pooled All
	USAtl	GoM	Cbn	WNA	RS	PG	WAus	EAus	IWP	Samples
n	41	58	24	123	38	11	43	53	96	272
nh	14	15	9	26	9	5	32	29	52	90
h	0.7939	0.7852	0.6159	0.7600	0.7240	0.7818	0.9712	0.9688	0.9735	0.8975
SD	0.0542	0.0482	0.1149	0.0373	0.0610	0.1073	0.0168	0.0092	0.0072	0.0156
π	0.0011	0.0012	0.0011	0.0012	0.0011	0.0019	0.0045	0.0044	0.0045	0.0108
SD	0.0008	0.0009	0.0008	0.0008	0.0008	0.0013	0.0025	0.0024	0.0024	0.0055

Table 2. *S. mokarran* summary statistics for each microsatellite locus by collection area. n = number of individuals, Na = number of alleles, A_r = allelic richness, GD = genetic diversity (Nei, 1987), H_E = expected frequency of heterozygotes, H_O = observed frequency of heterozygotes, Null = frequency of null alleles estimated using the Expectation Maximization (EM) algorithm of Dempster, Laird, and Rubin (1977). Far right column contains for each locus across all populations: (i) mean H_E, H_O values, (ii) overall A_r and F_{IS}, and (iii) total Na and null allele frequency. The bottom row contains for each population across all loci: (i) mean number of alleles, (ii) observed and expected frequency of heterozygotes, (iii) F_{IS}, (iv) mean null allele frequency. Last two lines of the bottom row, far right column are the overall F_{IS} and mean null allele frequency for all loci in all populations. Bold F_{IS} values are significant after sequential Bonferroni correction ($P < 0.05$), bold null allele values are >0.10. Collection areas are coded as in Table 1.

Table 2 (continued).

Locus	USAtl *n* = 39	GoM *n* = 54	Cbn *n* = 24	RS *n* = 36	PG *n* = 10	WAus *n* = 46	EAus *n* = 51	Mean/locus Total/locus
GhhA62								
Na	1	2	1	3	3	3	4	4
A_r	1.0000	1.1630	1.0000	2.7000	2.9930	2.8970	3.0360	2.6630
H_e	0.0000	0.0202	0.0000	0.4001	0.4650	0.5548	0.5323	0.2818
H_o	0.0000	0.0204	0.0000	0.3333	0.6000	0.4222	0.5490	0.2750
F_{IS}	NA	NA	NA	0.1805	-0.2414	0.2496	-0.0215	0.0990
Null	0.0010	0.0001	0.0010	0.0349	0.0000	0.0710	0.0000	0.0573
GhhC32								
Na	4	5	3	5	2	3	3	5
A_r	3.3150	3.4040	2.9340	3.0890	1.8000	2.7730	2.8450	3.7760
H_e	0.5890	0.6262	0.5444	0.3457	0.0950	0.4538	0.4462	0.4429
H_o	0.3947	0.6226	0.5217	0.3429	0.1000	0.4048	0.4510	0.4054
F_{IS}	**0.3416**	0.0152	0.0638	0.0228	NA	0.1199	-0.0009	0.0950
Null	0.1311	0.0070	0.0248	0.0000	0.0000	0.0551	0.0157	0.1235
GhhC7								
Na	4	4	4	4	4	6	6	9
A_r	3.3180	3.6990	3.3360	3.2010	3.6000	4.7950	4.0810	4.5120
H_e	0.6062	0.6452	0.5754	0.5531	0.5700	0.6741	0.6471	0.6102
H_o	0.6923	0.7115	0.5909	0.6000	0.8000	0.6744	0.6275	0.6709
F_{IS}	-0.1293	-0.0933	-0.0037	-0.0705	-0.3585	0.0114	0.0402	-0.0500
Null	0.0000	0.0000	0.0001	0.0000	0.0000	0.0191	0.0000	0.0218
GhhD9								
Na	3	3	2	2	2	2	2	3
A_r	2.2260	2.1450	1.9260	1.8030	2.0000	1.6200	1.8630	1.9470
H_e	0.2288	0.2298	0.2188	0.1567	0.3200	0.0997	0.1924	0.2066
H_o	0.1538	0.1481	0.2500	0.1714	0.4000	0.1053	0.1765	0.2007
F_{IS}	0.3391	**0.3634**	-0.1220	-0.0794	-0.2000	-0.0423	0.0927	0.1340
Null	0.1033	0.1064	0.0000	0.0000	0.0000	0.0000	0.0251	0.0519
GhhD103								
Na	12	16	12	20	12	21	21	30
A_r	7.0690	6.9740	7.9180	10.4400	12.0000	10.3230	10.0940	10.4270
H_e	0.8402	0.8255	0.8543	0.9208	0.8984	0.9167	0.9137	0.8814
H_o	0.8718	0.8367	0.7727	0.7714	0.8750	0.8333	0.9020	0.8376
F_{IS}	-0.0246	-0.0033	0.1185	0.1763	0.0926	0.1028	0.0227	0.0580
Null	0.0148	0.0000	0.0004	0.0819	0.0004	0.0441	0.0192	0.0449
Sle25								
Na	2	2	2	2	2	2	2	2
A_r	1.9500	1.9100	1.9540	1.5350	1.8000	1.3320	1.4040	1.7480
H_e	0.2604	0.2257	0.2491	0.0799	0.0950	0.0444	0.0571	0.1445
H_o	0.2564	0.2222	0.2917	0.0833	0.1000	0.0455	0.0588	0.1511
F_{IS}	0.0281	0.0245	-0.1500	-0.0294	NA	-0.0118	-0.0204	-0.0130
Null	0.0051	0.0051	0.0000	0.0000	0.0000	0.0000	0.0000	0.0050
Sle45								
Na	2	2	2	3	3	3	3	3
A_r	1.8230	1.6860	1.9540	2.6240	2.9960	2.7560	2.5400	2.3730
H_e	0.1672	0.1212	0.2491	0.3966	0.5150	0.4028	0.3547	0.3152
H_o	0.1842	0.1296	0.2917	0.3333	0.6000	0.2500	0.3922	0.3116
F_{IS}	-0.0882	-0.0600	-0.1500	0.1732	-0.1134	0.3913	-0.0959	0.0550
Null	0.0000	0.0000	0.0000	0.0743	0.0000	0.1250	0.0000	0.0339

Table 2 (continued).

Locus		USAtl $n = 39$	GoM $n = 54$	Cbn $n = 24$	RS $n = 36$	PG $n = 10$	WAus $n = 46$	EAus $n = 51$	Mean/locus Total/locus
Sle54									
	Na	4	5	3	5	4	6	5	6
	A_r	3.3500	3.3770	2.8140	4.2250	3.7680	4.4260	3.9810	4.0170
	H_e	0.6191	0.5837	0.5174	0.6069	0.5650	0.5782	0.5031	0.5676
	H_o	0.7368	0.5741	0.5833	0.6111	0.7000	0.5946	0.5098	0.6157
	F_{IS}	-0.1773	0.0258	-0.1065	0.0071	-0.1887	-0.0147	-0.0035	-0.0420
	Null	0.0000	0.0001	0.0192	0.0000	0.0000	0.0000	0.0000	0.0000
Sle71									
	Na	3	4	4	5	4	5	5	7
	A_r	2.9790	3.1430	3.3130	4.0010	3.8000	4.0350	3.8780	3.6610
	H_e	0.6361	0.6638	0.6085	0.6501	0.6650	0.7052	0.7026	0.6616
	H_o	0.6923	0.6296	0.6250	0.5556	0.6000	0.5952	0.5400	0.6054
	F_{IS}	-0.0755	0.0607	-0.0058	0.1592	0.1496	0.1677	0.2410	0.1080
	Null	0.0000	0.0213	0.0031	0.0399	0.0477	0.0458	0.0840	0.0348
Sle77									
	Na	21	24	17	29	16	32	35	43
	A_r	9.8070	9.5070	9.0530	11.9690	13.4740	12.0920	11.9110	11.3900
	H_e	0.9066	0.8838	0.8767	0.9387	0.9300	0.9459	0.9377	0.9171
	H_o	0.8718	0.8400	0.9167	0.9167	1.0000	0.9535	0.9412	0.9200
	F_{IS}	0.0514	0.0596	-0.0243	0.0375	-0.0227	0.0038	0.0062	0.0230
	Null	0.0000	0.0342	0.0000	0.0029	0.0000	0.0123	0.0000	0.0163
Sle81									
	Na	2	3	2	2	2	3	3	4
	A_r	2.0000	2.1510	2.0000	2.0000	2.0000	2.1770	2.1520	2.0910
	H_e	0.4997	0.5093	0.4922	0.4614	0.4200	0.3552	0.3798	0.4454
	H_o	0.3590	0.5094	0.4583	0.5000	0.2000	0.3571	0.3800	0.3948
	F_{IS}	0.2935	0.0092	0.0899	-0.0696	0.5610	0.0065	0.0096	0.0760
	Null	0.0939	0.0001	0.0229	0.0000	0.1649	0.0001	0.0002	0.0380
Sle86									
	Na	3	3	3	5	5	5	5	5
	A_r	2.9370	2.9690	2.9720	4.1400	4.7680	4.4280	4.4220	4.1290
	H_e	0.5135	0.5952	0.6033	0.7215	0.7250	0.7494	0.6952	0.6576
	H_o	0.5833	0.6226	0.7917	0.7647	0.7000	0.7500	0.6087	0.6887
	F_{IS}	-0.1221	-0.0365	-0.2929	-0.0451	0.0870	0.0118	0.1352	-0.0220
	Null	0.0000	0.0000	0.0000	0.0000	0.0000	0.0273	0.0409	0.0216
All loci									
	Na	5.0833	6.0000	4.5833	7.0000	4.9167	7.6667	7.8330	10.0833
	H_e	0.4889	0.4941	0.4824	0.5193	0.5220	0.5400	0.5301	0.5110
	H_o	0.4830	0.4889	0.5078	0.4986	0.5562	0.4988	0.5114	0.5064
	F_{IS}	0.0251	0.0202	-0.0307	0.0540	-0.0109	0.0885	0.0455	0.0360
	Null	0.0291	0.0145	0.0060	0.0195	0.0177	0.0333	0.0154	0.0194

Table 3. *S. mokarran* pairwise ϕST (mtCR) and FST (microsatellites).

Pairwise Φ_{ST} (mtCR)	WNA	RS	PG
WNA			
RS	0.02379 *		
PG	0.24376 **	0.17961 *	
IWP	0.87451 **	0.83971 **	0.81049 **

Pairwise F_{ST} (microsatellites)	WNA	NIO
WNA		
NIO	0.1161 **	
IWP	0.1113 **	0.01567 **

* indicates significance at $P < 0.02$

** indicates significance at $P < 0.000001$

Table 4. *S. mokarran* demographic statistics. Tajima's D statistic (*D*), Fu's F_S test (F_S), Fu & Li's *D**, Fu & Li's *F**, Harpending's Raggedness index (*Hri*), sum of squared differences from mismatch analyses (*SSD*), θ at time 0 (θ_0), θ at present time (θ_1), tau (τ), time since expansion in years before present (Time), and the migration parameter for the spatial expansion model (M). Values in bold indicate significance at $P < 0.025$ for Fu's F_S test, Fu and Li's *F**, and Fu & Li's *D**; and $P < 0.05$ for Tajima's *D*, *SSD*. Collection areas are coded as in Table 1.

Table 4a. Neutrality tests.

	Western North Atlantic (WNA)			Pooled	North Indian Ocean (NIO)		Indo West Pacific (IWP)		Pooled	Pooled All
	USAtl	GoM	Cbn	WNA	RS	PG	WAus	EAus	IWP	Samples
Tajima's D	**-1.7166**	-1.4007	-1.4720	**-1.8569**	-1.0630	0.8207	-0.8345	0.0575	-0.5400	1.1400
Fu's F_S	**-10.834**	**-10.072**	**-3.8549**	**-24.3134**	**-3.8365**	-0.3133	**-25.3751**	**-17.040**	**-25.5640**	**-23.9016**
Fu & Li's *D**	-2.1403	-1.6643	-1.0219	**-2.4319**	-0.0299	0.1491	-1.1760	-0.1029	-1.7936	-1.1687
Fu & Li's *F**	-2.3633	-1.9185	-1.2901	**-2.7094**	-0.4139	0.1393	-1.2554	-0.0555	-1.6185	-0.2652

Table 4b. Demographic expansion.

	Western North Atlantic (WNA)			Pooled	North Indian Ocean (NIO)		Indo West Pacific (IWP)		Pooled	Pooled All
	USAtl	GoM	Cbn	WNA	RS	PG	WAus	EAus	IWP	Samples
Hri	0.1250	0.0824	0.0510	0.0703	0.0551	0.1197	0.0081	0.0205	0.0120	0.0151
SSD	0.0141	0.0045	0.0065	0.0028	0.0007	0.0314	0.0007	0.0101	0.0044	0.0437
Θ_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9352	0.0000	0.0018	0.0000
Θ_1	99999	99999	1.9253	99999	99999	8.1863	26.5625	22.8125	21.7603	13.0859
τ	1.3809	1.4199	2.0547	1.3633	1.2695	3.1270	4.4258	5.9453	5.6777	24.4941
Time	105 375	108 355	156 795	104 033	96 879	238 620	337 735	453 692	433 272	1 869 168
τ min 95%	1.1037	0.6426	0.8066	0.9570	0.7695	0.6016	2.2188	2.5352	2.3594	0.6211
τ max 95%	1.8066	1.9023	3.9551	1.4648	1.7539	4.8281	6.5137	6.8106	6.1348	35.3496
Time 95% CI	84 222	49 036	61 555	73 032	58 723	45 906	169 315	193 460	180 046	47 396
Time 95% CI	137 866	145 169	301 815	111 783	133 842	368 438	497 064	519 719	468 149	2 697 558

Table 4c. Geographic expansion.

	Western North Atlantic (WNA)			Pooled	North Indian Ocean (NIO)		Indo West Pacific (IWP)		Pooled	Pooled All
	USAtl	GoM	Cbn	WNA	RS	PG	WAus	EAus	IWP	Samples
Hri	**0.1250**	0.0824	0.0510	0.0703	0.0551	0.1197	0.0081	0.0205	0.0120	0.0151
SSD	**0.0141**	0.0045	0.0028	0.0028	0.0007	0.0197	0.0016	0.0137	0.0075	0.0230
Θ	0.0007	0.0007	0.0007	0.0007	0.0026	0.0007	1.6781	0.7336	1.2021	4.8301
M	99999	99999	3.3884	99999	99999	6.0769	46.1491	34.2357	39.9001	1.0447
τ	1.3789	1.4200	1.6165	1.3619	1.2730	2.8985	3.5585	4.8928	4.1927	20.9051
Time	105 225	108 359	123 354	103 928	97 147	221 187	271 552	373 374	319 949	1 595 283
τ min 95%	0.7694	0.7150	0.1778	0.8154	0.6096	0.4164	2.5430	3.8201	2.2053	2.9314
τ max 95%	1.8927	1.7422	2.9445	1.5561	1.7226	4.4065	5.8211	6.5407	6.3257	28.3859
Time 95% CI	58 711	54 565	13 566	62 227	46 521	31 774	194 061	291 514	168 286	223 699
Time 95% CI	144 430	132 947	224 697	118 744	131 454	336 267	444 216	499 125	482 720	2 166 149

Table 5. *S. mokarran* demographic analyses. Collection localities are coded as in Table 1.

Table 5a. Mitochondrial control region. BEAST: Reported values inclue effective population size in millions (N_e) and time to most recent common ancestor in million years (TMRCA) with associated 95% highest posterior densities (HPD). ARLEQUIN: Time since population expansion in million years. LAMARC: Reported values include effective population in millions (N_e), exponential growth rate parameter (Population growth), and the migration rates in terms of $4N_em$. Migration rate 1 ➔ 2 indicates the rate of migration from population 1 into 2, and vice versa. Also reported are 95% confidence intervals.

	Beast Results				Arlequin	Lamarc Results							
Population or	N_e (mil.)		TMRCA (MY)		Time	N_e (million)		G		# migr per gen 1 ➔ 2		# migr per gen 2 ➔ 1	
Population Pair	mean	95% HPD	mean	95% HPD	(MY)	estimate	95% CI	estimate	95% CI	estimate	95% CI	estimate	95% CI
WNA	0.1188	0.01, 0.51	0.2488	0.12, 0.39	0.104	0.081	0.02, 1.00	5104	554, 12000				
RS	0.0620	0.00, 0.31	0.2005	0.09, 0.33	0.097	0.144	0.02, 2.87	7304	805, 14602				
PG	0.0382	0.00, 0.21	0.1985	0.08, 0.34	0.239	0.043	0.01, 2.40	4477	1, 12099				
Aus	0.8990	0.11, 3.73	0.7692	0.38, 1.22	0.433	0.413	0.16, 1.72	1302	453, 2813				
All	0.7548	0.10, 2.86	3.2485	1.58, 5.30	1.869								
WNA-RS										79.33	0.00, 359	30.97	0.00, 98.7
WNA-PG										6.02	0.00, 17.5	0	0.00, 5.61
WNA-Aus										0.29	0.00, 1.68	0.51	0.00, 2.00
RS-PG										13.94	0.00, 61.3	30.96	0.00, 159
RS-Aus										0.15	0.01, 5.94	0.61	0.14, 4.60
PG-Aus										0.68	0.00, 3.09	0.29	0.00, 1.07

Table 5b. Nuclear microsatellites. Mitochondrial control region. BEAST: Reported values inclue effective population size in millions (N_e) with associated 95% highest posterior densities (HPD). LAMARC: Reported values include effective population in millions (N_e), exponential growth rate parameter (Population growth), and the migration rates in terms of $4N_em$. Migration rate 1 ➔ 2 indicates the rate of migration from population 1 into 2, and vice versa. Also reported are 95% confidence intervals.

	Beast Results		Lamarc Results							
Population or	N_e (mil.)		N_e (million)		G		# migr per gen 1 ➔ 2		# migr per gen 2 ➔ 1	
Population Pair	mean	95% HPD	estimate	95% CI	estimate	95% CI	estimate	95% CI	estimate	95% CI
WNA	1.039	0.00, 2.82	0.133	0.06, 0.88	-0.004	-4.95, 1.46				
NIO	1.056	0.00, 5.12	1.12	0.80, 1.45	0.5	1.03, 1.98				
Aus	0.607	0.00, 3.22	0.789	0.51, 0.94	-0.11	-0.22, 1.48				
WNA-NIO							22.95	15.5, 34.2	0.95	0.56, 1.56
WNA-Aus							1.74	0.99, 2.61	0.37	0.26, 1.14
NIO-Aus							0.82	0.32, 1.50	14.9	11.1, 23.8

Table 5c. Nuclear microsatellite tests of demographic change. BOTTLENECK: Significance was estimated using the Wilcoxon sign rank deficiency test for heterozygosity excess. M-RATIO: M-values were estimated by M_P_VAL and the critical values were estimated using CRITICAL_M using a range of θ (0.01, 0.1, 1, 10). MSVAR: Reported values include the current population size (N0), the historic population size (N1), the time in million years since demographic change, and the associated 95% HPD.

	Bottleneck	Mratio	MSvar Results					
			N_0 (millions)		N_1 (millions)		Time (my)	
Population	Het. Excess	$M < M_c$	mode	95% HPD	mode	95% HPD	mode	95% HPD
WNA	significant	yes	0.052	0.00, 66.2	3.849	0.00, 8556	3.608	0.00, 7782
NIO	significant	yes	0.003	0.00, 19.7	4.226	0.00, 6380	0.094	0.00, 581.3
Aus	significant	yes	0.03	0.00, 66.2	4.326	0.00, 6380	1.104	0.00, 2426

Table 6. *S. mokarran* migrant and admixed individuals. For Structure analyses, if $q < 0.8$ for the sampling locality, the individual was assigned to the admixture class which had the highest posterior probability. The q-values for both populations are shown in parenthesis as (q-WAtl, q-NIO/IWP). Admixture classes are F1 (migrant parent) and F2 (migrant grandparent). For NEWHYBRIDS, individuals were assigned to a hybrid class if the posterior probability was greater than or equal to 0.75. Classes were pure western Atlantic (P-WAtl), pure NIO/IWP (P-IWP), F1 hybrid (F1), F2 hybrid (F2), backcross with western Atlantic (BC1), and backcross with NIO/IWP (BC2). Collection areas are coded as follows: (i) western Atlantic (WAtl): US Atlantic coast (USAtl), Gulf of Mexico (GoM), Caribbean (Cbn), and Western South Atlantic (WSA); (ii) Northern Indian Ocean (NIO): Red Sea (RS) and Persian Gulf (PG); (iii) Indo West Pacific (IWP): Indonesia (IND); western Australia (WAus); and eastern Australia (EAus).

			Migrant analysis	Admixture analyses (nuclear microsatellites)		
			(Nuclear microsatellites)		Admixture class	Admixture class
Specimen	Sampling locality	mtCR haplotype	GeneClass 1st generation migrants	Structure q-value (qWAtl, qNIO/IWP)	Structure (posterior prob)	NewHybrids (posterior prob)
OC-052	GoM	WAtl/NIO		WAtl (0.83. 0.17)	P-WAtl (0.85)	F2 (0.85)
OC-022	GoM	WAtl/NIO		WAtl (0.80, 0.20)	P-WAtl (0.90)	F2 (0.36)
OC-222	RS	WAtl/NIO	Migrant from WAtl	Admixed (0.45/0.55)	F2 (0.55)	F2 (0.81)
OC-080	IND	IWP		Admixed (0.36/0.64)	F2 (0.21)	F2 (0.95)
OC-090	RS	WAtl/NIO	Migrrant from WAtl	Admixed (0.24/0.76)	F1 (0.21), F2 (0.22)	F2 (0.59)
OC-214	RS	WAtl/NIO		NIO/IWP (0.13, 0.87)	P-Aus (0.75)	BC2 (0.74)
OC-263	WAus	IWP		NIO/IWP (0.10, 0.90)	P-Aus (0.67)	F2 (0.93)
OC-168	Eaus	IWP		NIO/IWP (0.05, 0.95)	P-Aus (0.87)	BC2 (0.77)
OC-175	Eaus	IWP		NIO/IWP (0.02, 0.98)	P-Aus (0.92)	BC2 (0.75)
OC-129	RS	WAtl/NIO	Migrant from IWP	NIO/IWP (0.02, 0.98)	P-Aus (0.99)	P-Aus (0.91)
OC-242	PG	WAtl/NIO	Migrant from IWP	NIO/IWP (0.01, 0.99)	P-Aus (0.99)	P-Aus (0.92)

Table 7. Geographic origin of market fins. The TCS and PCoA (mtCR) origin indicates which population the mitochondrial control region haplotype clusters with: the western North Atlantic / northern Indian Ocean (WNA/NIO) or Australia (Aus). The PCoA origin (msat) and STRUCTURE origin indicates which population the microsatellite genotype clusters with: the western North Atlantic (WNA) or the Australia / north Indian Ocean (Aus/NIO). The GeneClass origin indicates which of the three statistically differentiated populations had the highest posterior probability, western North Atlantic (WNA), northern Indian Ocean (NIO), or Australia (Aus). Posterior probabilities are shown in parenthesis, posterior probabilities greater than 0.50 are shown listed in descending order. All non-zero posterior probabilities are shown if all are less than 0.50.

Sample Number	Trader records	TCS Origin (mtCR)	PCoA Origin (mtCR)	PCoA Origin (msat)	Structure Origin (msat)	GeneClass Origin (msat)
F025	South America	WNA/NIO	WNA/NIO	Intermediate	Admixed	Aus (0.52)
F026	South America	Aus	Aus	Aus/NIO	Aus/NIO	Aus (0.72), NIO (0.57)
F027	South America	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.32), NIO (0.11), Aus (0.13)
F028	South America	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.98), Aus (0.57)
F032	Indian Ocean	Aus	Aus	Aus/NIO	Aus/NIO	NIO (0.82), Aus (0.61)
F034	South America	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.82)
F066	Unknown	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.86)
F074	Unknown	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.90)
F299	South East Asia	Aus	Aus	Aus/NIO	Aus/NIO	Aus (0.29)
F349	Unknown	Aus	Aus	Aus/NIO	Aus/NIO	Aus (0.61)
F370	South America	Aus	Aus	Aus/NIO	Aus/NIO	Aus (0.80)
F549	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.89), Aus (0.63)
F550	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.93)
F553	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.67)
F554	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.79)
F555	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.48), Aus (0.16)
F556	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.93)
F558	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	Aus (0.84), WNA (0.68), NIO (0.64)
F664	Brazil	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.96)
F720	Ecuador	WNA/NIO	WNA/NIO	WNA	WNA	WNA (0.49), Aus (0.30)

Figures

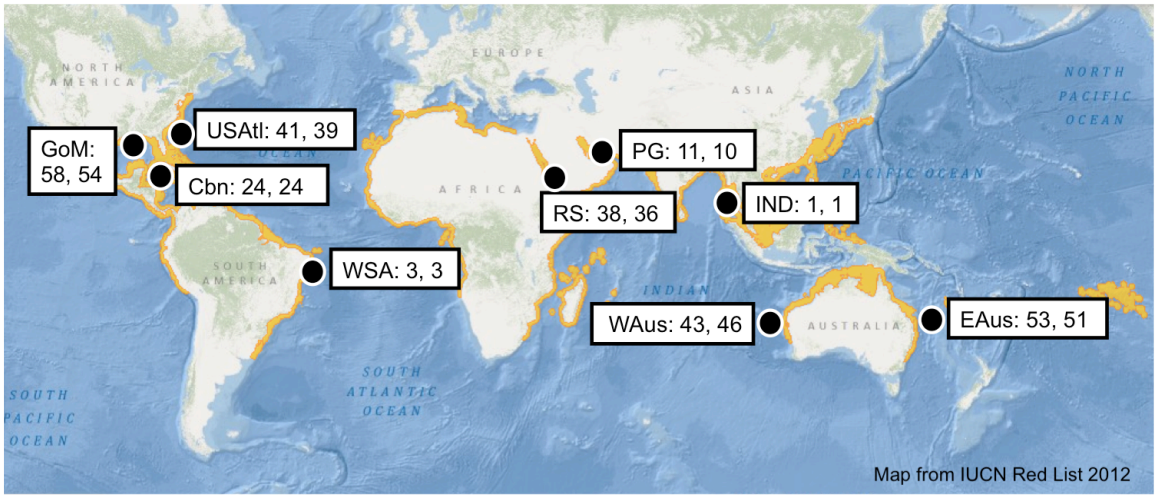


Figure 1. Map of *S. mokarran* distribution and sample sizes. Species distribution is shown in gold. Collection areas are represented by black circles, sample sizes are indicated in boxes, with the number of individuals sequenced (mtCR) shown first followed by the number of individuals genotyped with microsatellites. Collection areas are coded as follows: (i) western Atlantic (WAtl): US Atlantic coast (USAtl), Gulf of Mexico (GoM), Caribbean (Cbn), and Western South Atlantic (WSA); (ii) Northern Indian Ocean (NIO): Red Sea (RS) and Persian Gulf (PG); (iii) Indo West Pacific (IWP): Indonesia (IND); western Australia (WAus); and eastern Australia (EAus).

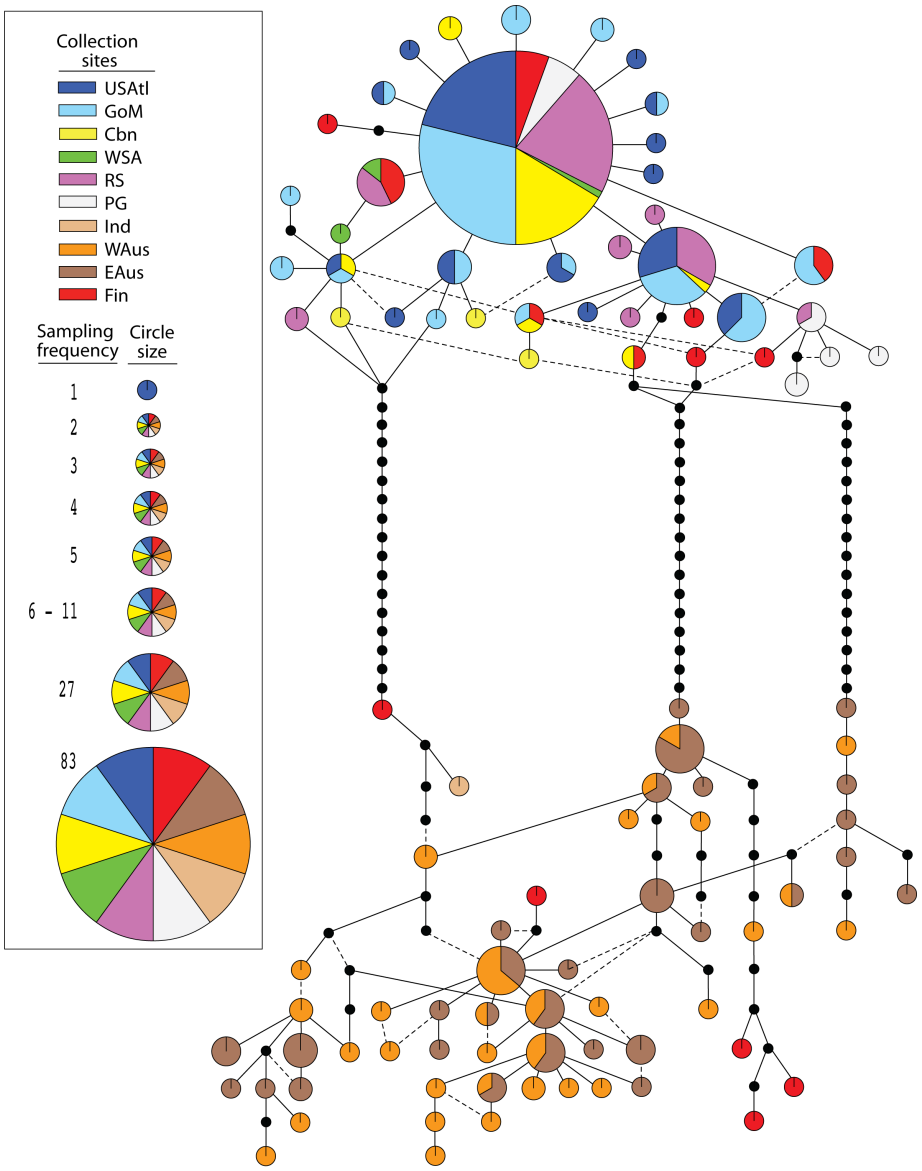


Figure 2. *S. mokarran* statistical parsimony haplotype network (mtCR). Circles represent individual haplotypes, with circle size proportional to sample frequency. Colored pie slices are proportional to the number of individuals from each sampling location with that haplotype. Unbroken connecting lines are equivalent to one mutation step and small black circles represent inferred, unsampled haplotypes. The WAtl/NIO clade at the top of the figure is separated from the Aus clade at the bottom by 18 mutation steps. Collection areas are coded as in Figure 1.

Figure 3a. Principal Coordinates (PCoA), mtCR

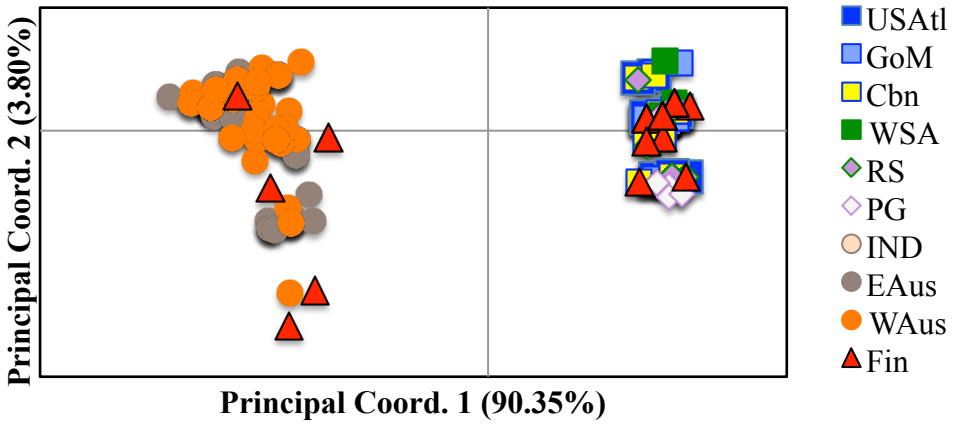


Figure 3b. Principal Coordinates (PCoA), microsatellites

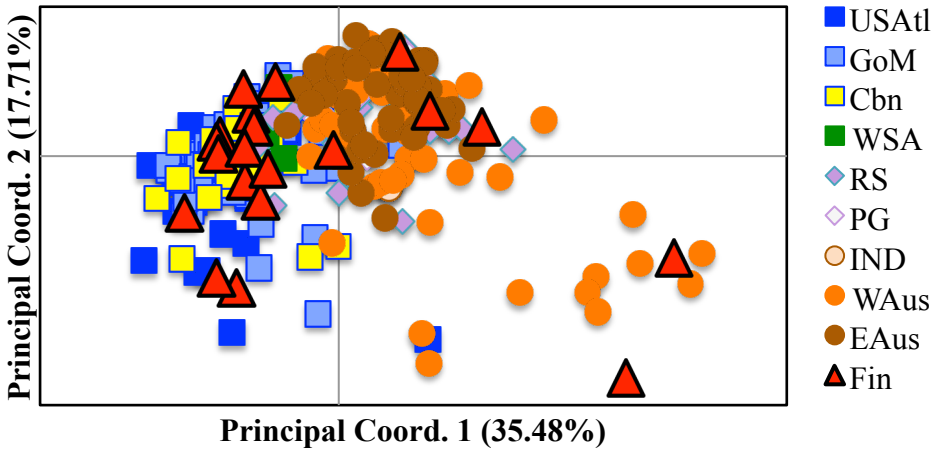


Figure 3. *S. mokarran* Principal Coordinate Analysis (PCoA) on (a) mitochondrial control region sequences and (b) nuclear microsatellite genotypes. Individuals from all collection areas were included, including where sample size was < 10. Colored shapes correspond to collection areas in the legends. The first two principal coordinate axes are shown with the amount of variance explained by each axis in parentheses. Collection areas are coded as in Figure 1.

Figure 4 *S. mokarran* STRUCTURE plot (microsatellites)

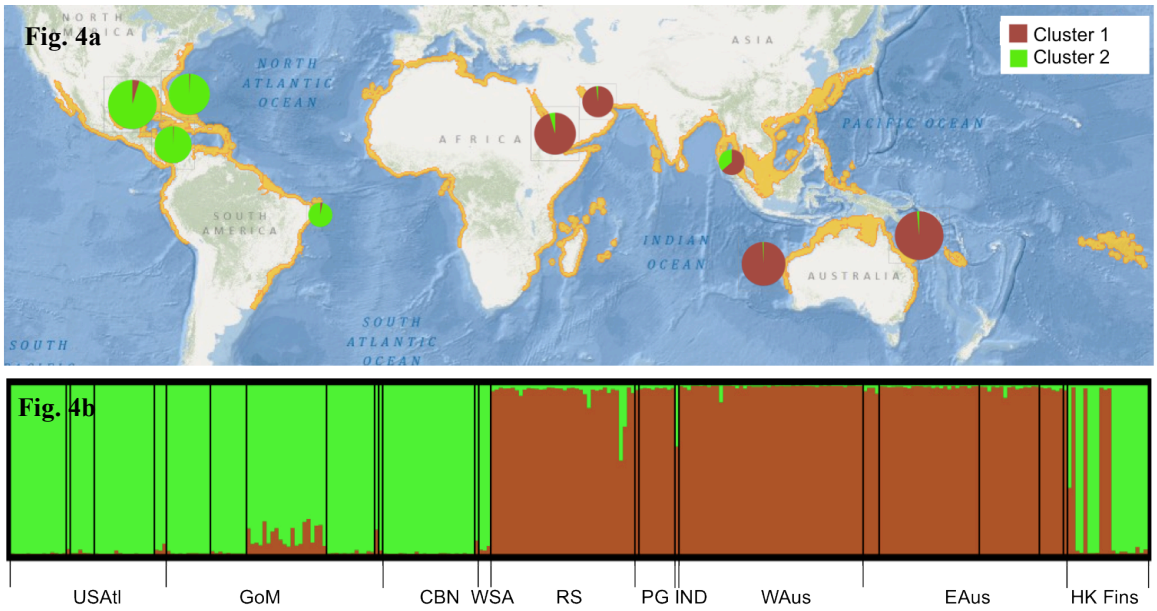
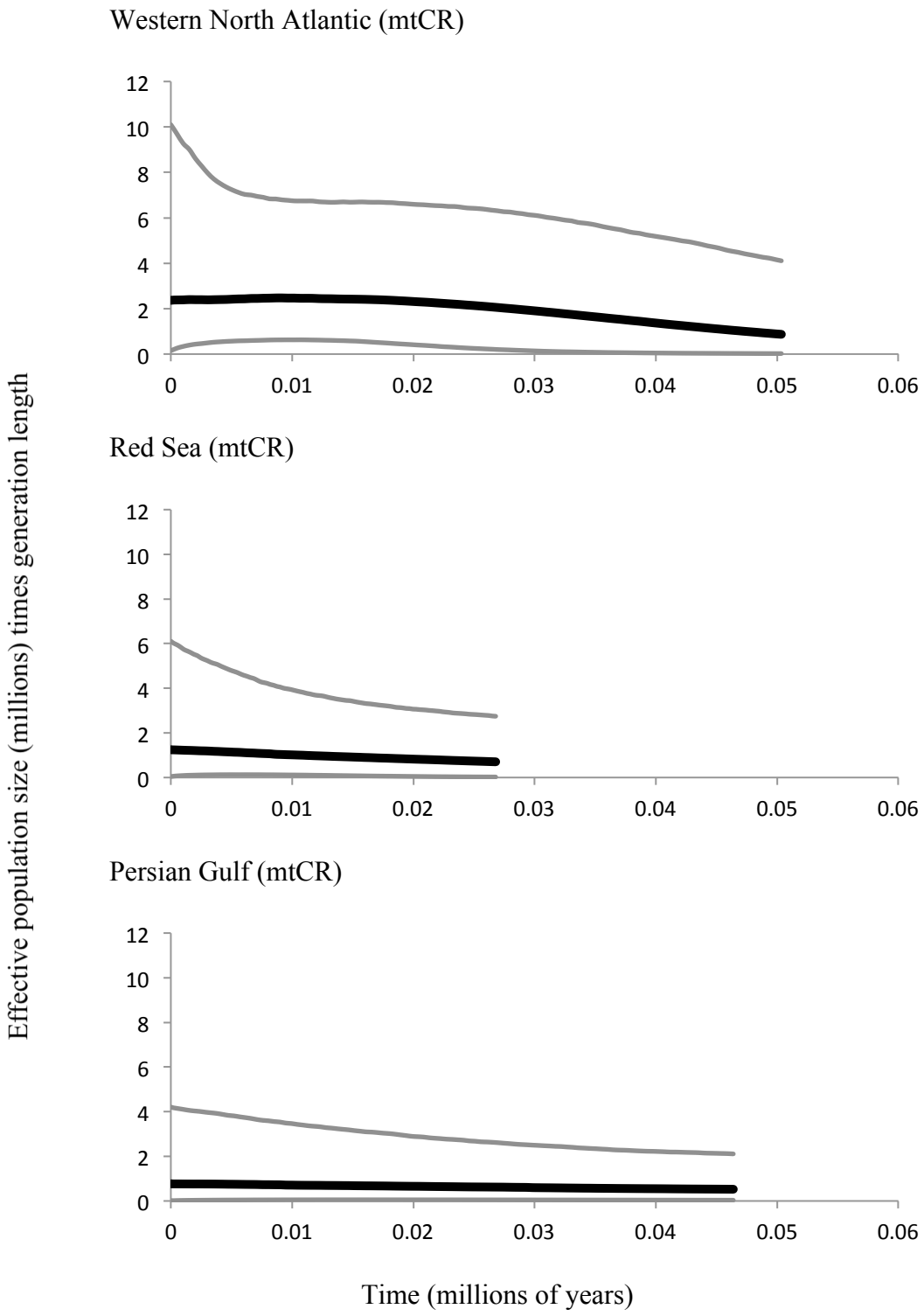


Figure 4. *S. mokarran* STRUCTURE analysis results (microsatellites). In Figure 4a, pie charts indicate the average proportional membership coefficient of individual sharks in the two populations inferred from nuclear microsatellite genotypes by the program STRUCTURE. Pie chart sizes are proportional to sample sizes. Figure 4b depicts assignment of individual sharks to each population in a conventional bar plot. Collection areas are coded as in Figure 1.

Figure 5.


Western North Atlantic (mtCR)
Red Sea (mtCR)
Persian Gulf (mtCR)
Effective population size (millions) times generation length
Time (millions of years)
0
2
4
6
8
10
12
0
0.01
0.02
0.03
0.04
0.05
0.06

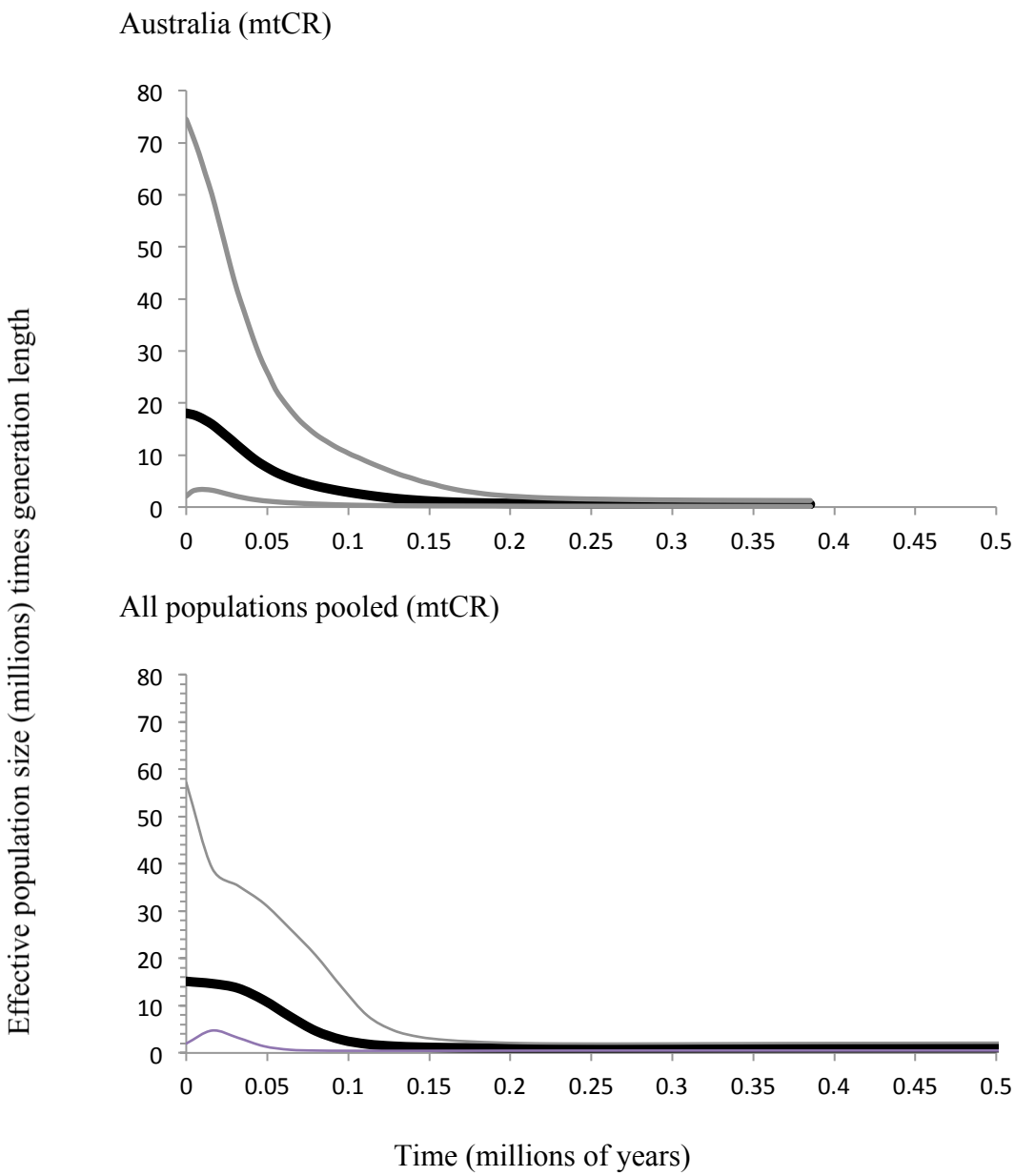


Figure 5. Bayesian (mtCR) skyline plots for each genetically differentiated *S. mokarran* population. Solid black line shows mean effective size through time. Grey lines show the 95% highest posterior density.

Supplementary Information.

Table S1. Microsatellite loci used to genotype *S. mokarran*.

Locus	Ta (^{o}C)	Source
Ghh-A6-2	55	Testerman *et al.* (In prep)
Ghh-C3-2	50	Testerman *et al.* (In prep)
Ghh-C7	50	Testerman *et al.* (In prep)
Ghh-D9	58	Testerman *et al.* (In prep)
Ghh-D10-3	52	Testerman *et al.* (In prep)
SLE025	50	Nance *et al.* (2009)
SLE045	58	Nance *et al.* (2009)
SLE054	58	Nance *et al.* (2009)
SLE071	50	Nance *et al.* (2009)
SLE077	58	Nance *et al.* (2009)
SLE081	58	Nance *et al.* (2009)
SLE086	58	Nance *et al.* (2009)

CHAPTER 4: Global population genetic structure, female philopatry and genetic connectivity in the smooth hammerhead shark (*Sphyrna zygaena*)

Abstract

The smooth hammerhead (*Sphyrna zygaena*) is a large, globally distributed, coastal-pelagic semi-oceanic shark typically found in temperate waters and occasionally in the tropics. It is caught in fisheries as a targeted species and as bycatch, and its fins are highly valued in the international fin trade. Despite its wide distribution and exploitation in fisheries, published information on smooth hammerhead biology and population dynamics is scarce. I evaluated the global population genetic structure of the smooth hammerhead using complete mitochondrial control region sequences (n=303, 1,090 bp) and 15 nuclear microsatellite loci (n=332) as markers. The mitochondrial and microsatellite data reveal strong genetic partitioning between the Atlantic and Indo-Pacific oceanic basins with no sharing of haplotypes or genotypes. Additionally, the mitochondrial data show genetic structuring within oceanic basins with at least 8 regional genetically distinct populations. The contrasting results from nuclear and mitochondrial markers may be indicative of female philopatry and male mediated gene flow in the smooth hammerhead, a species that reportedly uses coastal nursery areas. Statistical parsimony and coalescent analyses of mitochondrial DNA sequences reveal evolutionary relationships that are shallower than found in many other globally distributed sharks, and indicate recent genetic connectivity between the eastern South Pacific and western Atlantic. The high level of mitochondrial DNA-based population structure calls for smaller-scale regional management of the smooth hammerhead if its genetic legacy is to be conserved.

Introduction

The inherent dispersal potential of a species can impact its genetic population structure. For example, many marine species have high dispersal potential during larval stages or as highly vagile adults (Theisen *et al.* 2008; Díaz-Jaimes *et al.* 2010; Daly-Engel *et al.* 2012a) and expected low population structure, while others have low dispersal potential (Casado-Amezua *et al.* 2012; Botello *et al.* 2013) or exhibit philopatry to reproductive

areas or feeding grounds (Bowen & Karl 2007; Narum *et al.* 2007; Alter *et al.* 2009) that should result in increased population structure. Additionally, the marine environment presents many types of barriers to dispersal including continental obstructions, vast distances, varying depths and temperatures, food and habitat availability, and strong currents. All of these factors can impact the amount of population structure, genetic diversity, gene flow and demographic trends observed in the genetic signature of a given species, sometimes resulting in unexpected phylogenetic patterns (Rocha *et al.* 2007; DiBattista *et al.* 2013). Inferences based on genetic data can be crucial in understanding species demographic and biogeographic history, and in defining management units or genetically distinct lineages that are important for the conservation of exploited species.

Elasmobranchs have a long and rich evolutionary history that is echoed in their morphological and ecological diversity. They lack a larval phase, and juveniles frequently reside in distinct nursery areas during the first few years of life (Hueter *et al.* 2005; Wiley & Simpfendorfer 2007; Chapman *et al.* 2009; Tillett *et al.* 2012), so dispersal may occur at the subadult or adult stage. Some studies of elasmobranch population genetic structure have provided evidence of consistent patterns between adult vagility and population structure. For example angel sharks have highly restricted movement and high population structure (Gaida 1997), while circum-global movement patterns in the whale shark and basking shark correspond to low population structure (Hoelzel *et al.* 2006; Castro *et al.* 2007). Other biological characteristics and environmental factors can also impact genetic signatures. Biogeographic barriers have been implicated as potentially affecting genetic structure in zebra sharks (Dudgeon *et al.* 2009), Caribbean sharpnose sharks (Mendonça *et al.* 2011), spot tail sharks (Ovenden *et al.* 2009), great hammerhead sharks (Testerman, Chapter 3), and the white tip reef shark (Whitney *et al.* 2012). Female philopatry and male-mediated gene flow have been described for the great white shark (Pardini *et al.* 2001; Blower *et al.* 2012), the scalloped hammerhead shark (Duncan *et al.* 2006; Nance *et al.* 2011; Daly-Engel *et al.* 2012b), the black tip shark (Keeney *et al.* 2005; Sodré *et al.* 2012), and the blacktip reef shark (Mourier & Planes 2013). Thus, in elasmobranchs the observed patterns of genetic structure are not always as predicted by vagility alone.

The smooth hammerhead shark, *Sphyrna zygaena*, is the second largest hammerhead shark and is a strong, active swimmer. It has an amphitemperate distribution which is continuous around the southern coastlines of both South America and South Africa, although it is also found in sub-tropical and tropical waters (Compagno 1984; Last & Stevens 2009). It is commonly caught with a wide variety of gear by recreational fishers and in artisanal and commercial fisheries. Larger individuals are generally caught in pelagic fisheries, while juveniles are typically caught in near-shore fisheries as expected based on the species' reported use of shallow-water coastal nursery areas (Diemer *et al.* 2011). *S. zygaena* has a high post-capture mortality rate (Cortés *et al.* 2010; Coelho *et al.* 2012) and high-value fins (Last & Stevens 2009). Due to this susceptibility to exploitation, *S. zygaena* has been classified as vulnerable by the IUCN Red List (Casper *et al.* 2005) and received heightened conservation status via a recommended CITES Appendix II listing effective September 2014. Despite this level of conservation concern, little information has been published about this shark beyond gut content analyses (Smale & Cliff 1998; Last & Stevens 2009), and basic reproductive parameters such as a gestation period of 10-11 months with 20-50 pups per litter, an approximate size of 50 cm at birth, and maximum sizes of 370 - 400 cm (Compagno 1984; Last & Stevens 2009). Recently, Diemer *et al.* (2011) reported the results of a tag-recapture study that included 1 342 juvenile to adolescent *S. zygaena* along the coast of Southern Africa. Although most common in waters ranging from 16 - 22C, these young *S. zygaena* were also captured in colder waters ranging from 13 - 19C, which is similar to water temperature off the southern tip of South America. According to Diemer *et al.* (2011) the recapture rates and maximum movements they report (1.5% and 384 km) are probably lower than those reported for larger specimens in Kohler and Turner (2001) (3.2% and 1 122 km) due to life-stage related movement patterns.

Knowledge of genetic structure and patterns of gene flow can provide information to guide development of management strategies for this exploited species. Thus, I examined the amount of genetic variation based on the complete mitochondrial control region (mtCR) and 15 nuclear microsatellite loci from specimens of *S. zygaena* obtained throughout the species' global range. My study had three main objectives: (1) to investigate the extent of any genetic structure present in the species; (2) to estimate

historic and contemporary population sizes and levels of genetic connectivity; and (3) to evaluate patterns of genetic differentiation for concordance across mitochondrial and nuclear markers.

Materials and methods

Samples and DNA extraction

I obtained samples from other researchers, fisheries observer programs, and recreational fishers at 11 general collection areas in 3 ocean basins: (1) Atlantic Ocean (Atl): western North Atlantic (WNA – US Atlantic coast); western South Atlantic (WSA – Brazil); western Atlantic (WAtl): WNA and WSA combined; eastern North Atlantic (ENA – Azores); eastern Tropical Atlantic (ETA – Ivory Coast); (2) Indian Ocean (IO): western Indian Ocean (WIO – South Africa); eastern Indian Ocean (EIO – western Australia); and (3) Pacific Ocean: western South Pacific (WSP – eastern Australia and New Zealand); western North Pacific (WNP – Taiwan); eastern North Pacific (ENP – Baja California); eastern Tropical Pacific (ETP – Ecuador); eastern South Pacific (ESP – Chile). The final sample numbers used in analyses are shown in Figure 1. I also analyzed a single specimen obtained from a South African fin trader (FIN) that was caught in an unknown location. Tissue samples were obtained as fin clips or muscle punches from both fisheries and fisheries-independent sources and preserved in 95% ethanol. Genomic DNA was isolated from tissue samples using the QIAGEN DNeasy Tissue kit (QIAGEN, Inc., Valencia, CA).

DNA sequencing

The entire mitochondrial control region (mtCR) and some the flanking DNA from 338 *S. zygaena* individuals was amplified by primers CRF6 (5' AAG CGT CGA CCT TGT AAG TC 3') and DasR2 (5’ GCT GAA ACT TGC ATG TGT AA 3’). PCR reactions were performed in 50 µl volumes containing 1µl of unquantified genomic DNA, 200 µM of each dNTP, 12.5 pM each primer, 1 U HotStar *Taq*^{TM} DNA polymerase (QIAGEN, Inc.) and 5µl 10x *Taq*^{TM} reaction buffer. PCR cycling conditions were 95ºC for 15 minutes; followed by 35 cycles of 94ºC for 1 minute, 50ºC for 1 minute and 72ºC for 2 minutes; with a final extension of 5 minutes at 72ºC. In each set of PCR amplifications, a

negative control with no genomic DNA was included to check for contamination. Amplified products were purified using the QIAquick PCR Purification Kit (QIAGEN, Inc.) prior to direct cycle sequencing with a 16x dilution of BigDye 3.1 Terminator chemistry (Applied Biosystems, Inc., Foster City, CA) on both strands following the manufacturer's protocol. I used primers CRF6, DasR2, and two internal primers designed for this study, SzygCRF10 (5' ACA TCT CAT GTT CTG GTC AAG 3') and SzygCRR7 (5' GCT CAA GTT TAC CTG AAT GAA CCA G 3'). Sequencing reactions were purified using Dyex 2.0 Spin Kit (QIAGEN, Inc.) and sequenced on an AB3130 genetic analyzer (Applied Biosystems, Inc.). Sequences were aligned with GENEIOUS (Drummond *et al.* 2008) and alignments were checked and finalized by eye.

Microsatellite genotyping

I genotyped 405 *S. zygaena* individuals at 19 previously published nuclear microsatellite loci (Keeney *et al.* 2005; Nance *et al.* 2009; Testerman *et al.* In preparation). Annealing temperature and references for each locus are provided in Table S1. Microsatellite loci were amplified using forward primers with fluorescently labeled M13(-21) attached to their 5′ ends (Schuelke 2000). PCR reactions were performed in 12 μl volumes containing 1μl of unquantified genomic DNA, 200 μM of each dNTP, 2 pM forward primer, 5 pM reverse primer, 5 pM M13 tail, 1.5mM $MgCl_2$, 0.5 U HotStar *Taq*™ DNA polymerase (QIAGEN, Inc.) and 1.25μl 10x *Taq*™ reaction buffer. The few exceptions were 1.8 mM $MgCl_2$ for loci Ghh-F1, Sle81 and Tgr47, with locus Clim100 requiring 300 μM of each dNTP. PCR cycling conditions were 95ºC for 15 minutes; followed by 35 cycles of 94ºC for 1 minute, the locus specific annealing temperature for 1 minute, and 72ºC for 1 minute; with a final extension of 5 minutes at 72ºC. Fragments were separated on an AB 3130 genetic analyzer (Applied Biosystems, Inc.). Genotypes were scored using GENEMAPPER software version 3.7 (Applied Biosystems, Inc.) by comparison with the internal size standard LIZ 500 (Applied Biosystems, Inc.). Approximately 5% of the genotypes were reamplified and rescored to ensure genotyping repeatability and quality.

Individual samples that failed to amplify at more than 5 loci were excluded from further analyses. Genotypes were checked for duplicates, null alleles, scoring errors, and

large allele dropout using MICROCHECKER (Van Oosterhout *et al.* 2004) and MICROSATELLITE TOOLKIT (Park 2001) for EXCEL. Samples with identical multilocus genotypes or with up to two allele mismatches were considered to potentially be from the same individual, these genotypes were re-evaluated and corrected if appropriate. Several juveniles and individuals of unknown age were sampled at various collection areas. Because juvenile sharks may use the same nursery site for a few years (Wiley & Simpfendorfer 2007; Chapman *et al.* 2009), it is possible that these juveniles may have been either full or half siblings, and I used the pairwise hypothesis testing option in ML-RELATE (Kalinowski *et al.* 2006) to identify related individuals (full sibs, half sibs, or parent-offspring) in each collection area. Likelihood ratios of 10 000 random dyads were simulated, and the hypothesis that a pair was related was accepted when the probability of their likelihood ratio was <0.01. When duplicate or related samples were found, one individual in each pair was randomly selected for inclusion in further analyses of mtDNA sequences and nuclear genotypes.

Statistical analysis

Due to small sample sizes from the eastern North Atlantic, eastern Tropical Atlantic, and eastern Indian Ocean, data from those collection areas were not used in statistical or population-level analyses but were included in individual level analyses. For all statistical analyses with multiple tests, I evaluated significance levels after sequential Bonferroni correction (Rice 1989).

mtCR DNA sequence analysis – genetic diversity and population structure

The ARLEQUIN software v3.5.1.2 (Excoffier *et al.* 2005) was used to calculate the number of haplotypes (*nh*), haplotype diversity (*h*) and nucleotide diversity (π) for each collection area with n > 10 samples. Genetic structure within a geographic context was investigated using three different approaches. First, the amount of genetic differentiation within and among populations was examined by an analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) with the Tamura & Nei (Tamura & Nei 1993; Bowen *et al.* 2005) distance as implemented in ARLEQUIN. AMOVA calculates analogs to Wright's F-statistics (Wright 1951, 1965), designated Φ_{ST}, based on the allelic content of haplotypes and haplotype frequencies. I initially calculated pairwise Φ_{ST} values among the eight

collection areas with n > 10 samples. Neighboring collection areas that had non-significant pairwise Φ_{ST}s ($p > 0.05$) were pooled into putative populations for subsequent population level and statistical analyses. The amount of genetic differentiation within and among these populations was evaluated through hierarchical AMOVA. The significance of the covariance components was tested using non-parametric permutation procedures with 1 000 permutations. Secondly, SAMOVA analyses of pairwise differentiation (Dupanloup *et al.* 2002) were conducted with a range of K (2-12) and 1000 permutations. The most likely number of genetically differentiated groups was assumed when F_{CT} was maximized. Finally, MEGA version 3.1 (Kumar *et al.* 2004) was used to calculate the pairwise genetic distances (d) and within group means (π) within and among genetically distinct populations under the Tamura Nei evolution model. The variance within populations was corrected by d corr = $d - ((\pi_1 - \pi_2)/2)$. Genetic distances between Sphyrnid species from (Testerman, Chapter 1) were included for comparison.

Common haplotypes were identified using DNACOLLAPSER v1.0 (Fredsted 2006). Two methods were used to visualize patterns of differentiation among individuals and determine the presence of distinct genetic clusters. First, the evolutionary relationships between haplotypes were evaluated in unrooted parsimony-based haplotype networks constructed using the Templeton *et al.* (1992) method as implemented in TCS v1.21 (Clement *et al.* 2000) with the connection limit set initially at 95%. Ambiguities in the networks were resolved using criteria based on coalescent theory (Crandall & Templeton 1993; Pfenninger & Posada 2002; Richards *et al.* 2009); alternate connections are shown as dashed lines. Secondly, principal coordinate analysis (PCoA) as implemented in GENEALEX v6.5b3 (Peakall & Smouse 2006; Peakall & Smouse 2012) was used to examine the amount of genetic differentiation. Eigenvectors were calculated from a covariance matrix and the first two coordinates were plotted.

mtCR DNA sequence analysis – population demographics

A variety of approaches were used to evaluate population demographics, but generation time and mutation rates are required to translate parameter estimates into demographic terms. Because this information is lacking for *S. zygaena*, the 20-year generation time and mtCR mutation rate of 6.05×10^{-9} substitutions per site per year estimated for its congener *Sphyrna lewini* by Nance *et al.* (2011) were used in demographic calculations.

Demographic summary statistics were calculated in ARLEQUIN. Tajima's D statistic (Aris-Brosou & Excoffier 1996; Tajima 1996) tests the hypothesis of selective neutrality and population equilibrium; however significant *D* values can also result from factors such as heterogeneity of mutation rates. Fu's F_S (Fu 1997) evaluates neutrality of mutations, and is very sensitive to population expansion which leads to large negative values of F_S. For comparison purposes, the D* and F* statistics of Fu and Li (1993) were calculated in DNASP v5.1 (Librado & Rozas 2009). Population expansion is suggested when Fu's F_S is significant and D* and F* statistics are not whereas the reverse combination suggests selection (Fu 1997). The mismatch distribution analysis (Schneider & Excoffier 1999) in ARLEQUIN computes the number of differences between pairs of haplotypes and uses a non-linear least-square approach to estimate parameters of a sudden demographic (Rogers & Harpending 1992) and geographic (Ray *et al.* 2003) expansion. The sum of squares deviations (SSD) between the observed and expected mismatch distribution tests the validity of the expansion model. Harpending's raggedness index (*Hri*) (Harpending 1994) tests the fit of the mismatch distribution and provides indications of population expansion, with larger values of *Hri* typical of multimodal distributions found in stationary populations and smaller, non-significant values for unimodal or smoother distributions typical of expanding populations. Tau (τ) is calculated in the mismatch analysis and is a relative measure of the time in generations since population expansion ($\tau = 2\mu t$). Time since expansion in years (Time) was calculated by $t = \tau/2\mu$ where μ is the mutation rate per site per generation. For the geographic expansion model, *m* is the migration rate between the sampled deme and a population of infinite size after T generations, $M = 2N_e m$ and N_e is the effective population size.

MDIV (Nielsen & Wakeley 2001), a Markov chain Monte Carlo (MCMC) method, was used to estimate the scaled parameters Θ ($\Theta = N_{ef}\mu$), M ($m = \Theta*M/2$) and T ($t = 2TN_{ef}$) between each pair of genetically differentiated populations identified by pairwise Φ_{ST} analysis. The HKY model of evolution was used because, according to the software authors, it is a more accurate model of DNA evolution although it is computationally slower than the other option available in the program. Chain length was 5×10^6 cycles and burn-in time was set to 10% of the total length of cycles. For the

initial data collection runs, Mmax and Tmax were set at 100 and 20, respectively, to determine appropriate upper bounds and revised for the final data collection runs. For each population comparison, the data set was analyzed at least three times with different random seeds to evaluate convergence. For each of the output parameters, the value with the highest likelihood was selected as the best estimate. The MDIV analyses were run using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation.

The Bayesian implementation in LAMARC v2.1.8 (Kuhner 2006), a coalescent-based MCMC method, was used to estimate Θ ($\Theta = N_e\mu$) and growth rates for genetically differentiated populations. The parameter N_e is reported here to facilitate comparison with N_e derived from nuclear microsatellite data, female effective population sizes can be calculated by $\Theta = 2N_{ef}\mu$. I report directional migration rates between population pairs in terms of the biologically significant parameter $4N_em$ ($4N_em = M\Theta_{rec}$ where M is the LAMARC migration output parameter and Θ_{rec} is Θ of the receiving population) because the force of migration becomes strong enough to overcome genetic drift when 4Nem > 1 (Kuhner 2006). Initial data runs were conducted with wide priors to determine appropriate upper bounds and the priors were revised for final data runs. Each data set was analyzed at least three times with different random seeds and run length of 1.0×10^7 generations, while the final input prior values and all other inputs were held constant to evaluate convergence. For each of the output parameters, the means from each of the final runs were averaged to obtain the best estimate.

Changes in effective population size over time were inferred through extended Bayesian skyline plots (EBSP) as implemented in BEAST v1.7.5 (Drummond *et al.* 2012). The HKY+I+G model was used because it most closely approximated the best fit model of DNA evolution estimated using the AICc in ModelTest v3.7 (Posada & Crandall 1998). Each population was analyzed three times using a constant skyline model with 10 groups and automatic optimization of operators. Genealogies and parameters were sampled every 1,000 iterations. The analyses were run for 2×10^7 iterations with a strict molecular clock set at the *S. lewini* mutation rate (Nance *et al.* 2011). Convergence was assumed when the ESS for each parameter in each replicate run exceeded 200. Skyline plots were generated in TRACER v1.5 (Drummond & Rambaut 2007).

Microsatellite analysis

Input files were prepared for microsatellite data analyses using the software CREATE v. 1.36 (Coombs *et al.* 2008). Microsatellite loci were checked for deviations from Hardy-Weinberg equilibrium and for evidence of linkage disequilibrium using GENEPOP v4.1 (Rousset 2008). Exact tests of Hardy-Weinberg equilibrium were performed using the Markov chain method (Guo & Thompson 1992; Rousset & Raymond 1995), with 10 000 dememorizations, 20 batches and 5 000 iterations per batch. Global tests across loci and collection areas used Fisher's method. (Guo & Thompson 1992; Rousset & Raymond 1995).

Microsatellites – genetic diversity and population structure

The software GENETIX v4.05.2 (Belkhir *et al.* 1996) was used to estimate observed (H_o) and expected (H_e) levels of heterozygosity, number of alleles, and inbreeding coefficients (F_{IS}). FSTAT v2.9.3.2 (Goudet 1995, 2001) was used to calculate genetic diversity (Nei 1987) and allelic richness (A_r) (El Mousadik & Petit 1996). The frequency of null alleles was estimated using the Expectation Maximization (EM) algorithm of (Dempster *et al.* 1977) as implemented in GENEPOP.

GENEPOP was used to test for genic differentiation among the eight collection areas with n > 10 samples. Fisher's exact tests for significance of genic differentiation were performed using the Markov chain method, with 10 000 dememorizations, 20 batches and 5 000 iterations per batch. Global tests across loci and collection areas used Fisher's method. Neighboring collection areas that were not significantly differentiated from each other were pooled into putative populations for subsequent statistical and population level analyses.

Genetic population structure was assessed using three different approaches. First, individuals were assigned to populations using the Bayesian clustering algorithm in STRUCTURE version 2.3.4 (Pritchard et al. 2000; Falush et al. 2003) using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation. I ran STRUCTURE with the collection areas as prior information, the admixture model, correlated allele frequencies, and a burn-in period of 100 000 MCMC generations followed by 200 000 iterations. I varied the number of populations from k = 1 through k = 10 with 10 replicates for each k. STRUCTURE

HARVESTER v0.6.93 (Earl & vonHoldt 2012) was used to determine the most likely number of distinct genetic clusters by evaluating the logarithm of the probability of the data (lnP(D | K)) (Pritchard *et al.* 2000) and estimates of Δk (Evanno *et al.* 2005). Each individual's admixture proportions were averaged over the 10 replicates for the most likely k using the program CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007) and the output was graphically displayed by the program DISTRUCT v1.1 (Rosenberg 2004). I used a second non-MCMC algorithm that uses an iterative reallocation method to assign individuals into k distinct populations without any *a priori* sample location information. FLOCK v2.0 (Duchesne & Turgeon 2009; Duchesne & Turgeon 2012) was run starting with $k = 2$ and increasing until one stopping condition was reached, with 50 runs per k and 20 iterations per run. Both programs were used because STRUCTURE may have more power when migration rates are low while FLOCK may have more power under conditions of high, sustained migration (Duchesne & Turgeon 2012). Finally, principal coordinate analyses (PCoA) was performed in GENEALEX to visualize genetic differentiation among individuals as described above.

Microsatellites – population demographics

Population demographics for the genetically differentiated populations identified by pairwise F_{ST} were assessed using both statistical and coalescent-based approaches. First I tested for evidence of bottlenecks in each of the genetically distinct populations using two statistical moment-based estimators, the sign-rank test and the M-ratio test. BOTTLENECK v1.2.02 (Piry *et al.* 1999) calculates statistics that test for departures from equilibrium patterns of heterozygosity that can be disrupted when the effective population size is significantly reduced. The Wilcoxon sign-rank test was used because it is the most powerful when fewer than 20 loci are analyzed (Piry *et al.* 1999). Calculations were performed assuming three different microsatellite mutation models: the infinite alleles model (IAM), the single-step model (SSM), and the two phase model (TPM). For the TPM, variance was set to 12 and 95% single step mutations were selected as recommended by the authors. Significance was assessed over 10 000 replicates.

To further test for population declines, I calculated the *M*-ratio statistic using the software *M_P_val* (Garza & Williamson 2001) and compared it to a simulated

equilibrium distribution calculated in *critical_M* selecting the TPM, 10 000 replicates, with the percentage of mutations that follow the single step model (p_s=0.9) and the mean size of larger mutations (Δ_g=3.5) as suggested by the software authors. Critical values (M_C) were calculated for a range of θ (0.01, 0.1, 1 and 10), and statistical significance was assessed over 10 000 replicates.

Changes in effective population size were assessed using three coalescent-based methods. First, the hierarchical Bayesian MCMC model implemented in MSVAR 1.3 (Storz & Beaumont 2002) was used to estimate parameters including the current population size (N_0), the ancestral population size (N_1), the time of population size change (t), and the mutation rate (μ). This estimated microsatellite mutation rate was used in further demographic analyses. MSVAR parameters included the exponential growth model, 5 replicate runs with varied priors and broad hyperpriors for the model parameters, and run lengths of 9 x 10^8 or 5 x 10^8 steps with output reported every 20 000 steps. After removing a 10% burn-in, convergence was assessed by calculating the Gelman-Rubin multivariate scale reduction factor (PRSF) and the effective sample size (ESS) across the 5 independent runs using the CODA package in the software R v2.15.1 (R Development Core Team 2012). Estimates of the mode and 95% highest posterior density (HPD) intervals were obtained using the modeest v1.14 and coda v0.15-2 packages, respectively, in R.

The Bayesian implementation in LAMARC v2.1.8 (Kuhner 2006) was also employed to estimate population Θ ($\Theta=4N_e\mu$) and growth rates for genetically differentiated populations, and per-generation migration rates ($m=M*\Theta_{rec}$) between these populations. The search strategy was similar to that used for the mtCR analyses, however, each data set was run for one extremely long final run of 2 x 10^7 generations.

Finally, I inferred changes in effective population size over time as well as a point estimate of the current effective population size through extended Bayesian skyline plots (EBSP) as implemented in BEAST v1.7.3 (Drummond *et al.* 2012). Three separate analyses were run for each population using the equal rate proportionality, unbiased mutational estimate, single phase, and a strict molecular clock with broad uniform priors. The analyses were run for the maximum of 2.147 x 10^9 iterations. Convergence was assumed when the ESS for each parameter in each run exceeded 200.

Microsatellites – Individual assignment

The number of first generation migrants was analyzed using the Bayesian individual assignment method as implemented in GENECLASS 2.0 (Piry *et al.* 2004) at $p < 0.05$. Selected parameters included L=L_Home as not all possible source populations were sampled, the Rannala and Mountain (1997) Bayesian criteria, and Monte-Carlo resampling (Paetkau *et al.* 2004) with 10 000 simulated individuals ($\alpha = 0.05$) to assess probabilities. The number of first, second and third generation migrants were also assessed using the prior population information and Gensback = 2 options in STRUCTURE with the run parameters described above, except the number of populations was fixed at 2 based on the Atlantic and Indo-Pacific populations defined by F_{ST} analyses. In these analyses, STRUCTURE assesses the proportion of ancestry (q) in both possible populations (WAtl or NIO/IWP). Individuals that are not migrants or admixed should have a posterior probability of being from the assumed population that is close to 1.0, with lower values indicating a migrant or admixed individual. A cut-off value of 0.8 was chosen such that if q was < 0.8 for the sampling locality, that individual was assigned to the class with the highest posterior probability. Individuals were assigned as either a migrant or as having at least one migrant parent (F1) or grandparent (F2). To determine the potential source population of the single market fin, GENECLASS was used to calculate the probability that the Atlantic and Indo-Pacific were possible source populations using 10 000 simulated individuals, and a probability cut-off of 0.5. I also evaluated the q-value of this genotype from STRUCTURE and included the mtCR sequence in the TCS haplotype network.

Mito-nuclear discordance

Possible sources of mito-nuclear discordance were investigated by transformation of the mtCR Φ_{ST} and standardization of nuclear F_{ST} values to facilitate direct comparisons. The mitochondrial Φ_{ST} values calculated above measure genetic differentiation based on both haplotype frequency and the genetic distance between haplotypes, while nuclear F_{ST} measures genetic differentiation based upon allele frequency shifts between locations and does not consider genetic distance. Additionally, there is a strong downward bias inherent in F_{ST} estimated from highly polymorphic loci, confounding comparisons of F_{ST} analogues from markers with different mutation rates such as mtCR and microsatellites.

To correct for these biases, an analogue of mitochondrial F_{ST} (mtF_{ST}) was calculated in ARLEQUIN based solely upon haplotype frequencies. A standardized nuclear F'_{ST} (Hedrick 2005) was then calculated using RECODEDATA (Meirmans 2006) which is based upon a recalculated maximum possible value and alleviates the dependence on within-population genetic variation.

Results

Genetic variation and geographic structure based on mitochondrial DNA sequence

After deleting 4 duplicates and 31 related samples as described above, sequence data from the complete mtCR sequence (1 090 nucleotides) of 303 individuals identified 19 polymorphic sites that defined 31 haplotypes. Overall haplotype and nucleotide diversity were high at 0.8841 and 0.0032, respectively. Collection areas in the Indo-Pacific generally have higher haplotype and nucleotide diversities, while collection areas in the Atlantic have somewhat lower levels of diversity (Table 1).

Pairwise Φ_{ST} values between all eight collection areas with n > 10 were highly significant (Table 2a). In general, intra-basin pairwise Φ_{ST} values were smaller than those from inter-basin comparisons. Heirarchical AMOVA (Table 2b) revealed significant partitioning between the Atlantic and Indo-Pacific basins as well as among collection areas. Results from SAMOVA (Table 2c) were concordant, with maximum F_{CT} occurring for K = 8. The net pairwise genetic distances (Table 3) between *S. zygaena* populations ranged from 0.00004 to 0.0050 while inter-specific distances ranged from 0.1382 to 0.1967.

A geographic pattern of shallow genetic differentiation is present in the statistical parsimony haplotype network in Figure 2. There is no sharing of haplotypes between the Atlantic and Indo-Pacific basins. Four distinct clades are indicated by shading. One clade is comprised of haplotypes from the Atlantic Ocean, one of haplotypes from the eastern Tropical/South Pacific, one of haplotypes primarily from the eastern North Pacific, and one is comprised of individuals from the western Indian Ocean. Each of these clades contains a single high frequency haplotype with other lower frequency haplotypes. There is a central mixed clade of haplotypes from Indo-Pacific collection areas that separates the southwest Indian Ocean and eastern North Pacific clades from the

eastern Tropical/South Pacific and Atlantic clades. The Atlantic and the eastern Tropical/South Pacific clades are separated by two mutational steps, while the majority of haplotypes are separated by a single mutational step.

Principal coordinate analysis provided clear evidence of geographic segregation (Figure 3a). Individuals from the Atlantic were located in the upper right quadrant, most individuals from the eastern Tropical and South Pacific and the western North Pacific were in the two right quadrants, most individuals from the western Indian Ocean were in the two left quadrants, eastern north pacific individuals were in the two left quadrants, and most western south pacific individuals were in the lower left quadrant.

Genetic variation and geographic structure based on nuclear microsatellites

A total of 29 samples amplified at less than 75% of all loci. Four loci (GhhG5, GhhF1, Sle038, and Sle081) were removed from the data set due to patterns of null alleles or linkage disequilibrium across collection areas. Analysis of the remaining 15 loci revealed four pairs of duplicate multilocus genotypes, 37 pairs of probable full siblings, and three pairs of probable half siblings, with both members of each pair caught in the same collection area. After removing one random genotype from each of these pairs, the multilocus genotypes were determined for a total of 332 samples at 15 microsatellite loci. Polymorphism ranged from 4 to 56 alleles per locus (Table 4). There was no evidence of scoring errors or large allele dropout. A few loci exhibited lower than expected levels of heterozygosity in some collection areas, but there was no systematic pattern indicative of null alleles and no loci exhibited a null allele frequency > 0.10 in collection areas with n > 20. Similarly, a few collection areas showed evidence of linkage disequilibrium at some pairs of loci, but none of the loci showed evidence of linkage disequilibrium consistently across sampling locations. One locus (GhhD1) had significant F_{IS} values in the western North Atlantic and the western Indian Ocean, but again none of the loci had significant F_{IS} values consistently across sampling locations. Although diversity indices varied across loci, most had slightly higher levels of allelic richness and observed heterozygosity in the Indo-Pacific than in the Atlantic.

Pairwise F_{ST} values were small and non-significant among collection areas within the Atlantic and within the Indo-Pacific while pairwise F_{ST} values were large and significant for inter-basin comparisons (Table 5). Collection areas in the Atlantic and

those in the Indo-Pacific were considered to comprise two genetically distinct populations for further analyses. PCoA analysis (Figure 3b) indicates the presence of two overlapping clusters, one in the Atlantic and one comprising the Indo-Pacific individuals. Bayesian STRUCTURE analyses (Figure 4) and the FLOCK analyses also indicate the presence of the same k = 2 genetically distinct clusters. Separate STRUCTURE and FLOCK analyses (data not shown) of only the Indo-Pacific and only the Atlantic samples both returned k = 1.

Mito-nuclear discordance

Pairwise F_{ST} statistics are presented in Table 5. The mtF_{ST} is generally smaller than the $mt\Phi_{ST}$ as expected, and all values remain significant. Mierman's standardized F'_{ST} values are larger than traditional F_{ST} values, although they retain the same patterns and significance estimates. The pairwise F'_{ST} value between the Atlantic and Indo-Pacific was large (0.3593) and significant ($P < 0.00001$).

Mitochondrial demographic analyses

Tajima's *D*, Fu's F_S, and Fu and Li's D* and F* are generally small but non-significant (Table 6a), which does not suggest that either expansion or selection has occurred. However, the sum of squares deviations from the mismatch analyses cannot reject a sudden expansion model and Harpending's raggedness index suggests unimodal distributions indicative of population expansion, although these analyses do not distinguish between demographic (Table 6b) or geographic (Table 6c) expansion. Time since expansion varies by collection area with no clear trends evident, although time since expansion is greater for the pooled Indo-Pacific than for the Atlantic samples. Estimates of divergence times obtained in MDIV (Table S2) ranged from 5 - 65 kya, although the divergence time estimate for the eastern Tropical and South Pacific sites was less than 0.5 kya. Biologically significant migration rates ($4N_em$) estimated in LAMARC and MDIV are depicted in Figure 5 with details of pairwise comparisons in Table S2. Generally migration rates were biologically significant (i) among the western Indian Ocean, the western North Pacific and the eastern North Pacific; (ii) from the western Indian Ocean and western North Pacific to the eastern South Pacific; (iii) from the western South Pacific to the western Indian Ocean, eastern North Pacific and eastern

South Pacific; (iv) between the eastern Tropical and South Pacific; and (v) from the western North and South Atlantic to the eastern South Pacific. The confidence intervals for growth rates estimated in LAMARC (Table S3) overlapped zero for the western North Atlantic, western South Atlantic, western South Pacific, western North Pacific, and eastern North Pacific indicating possible population declines in these regions. Estimates of N_{ef} from LAMARC (Figure 6a and Table S3) were highest for the western Indian Ocean and eastern South Pacific, intermediate for the western North Pacific and eastern North and Tropical Pacific, slightly lower for the western South Atlantic and western South Pacific, and lowest for the western North and South Atlantic. The patterns were similar but not as distinct in estimates of N_{ef} obtained from BEAST (Figure 6b and Table S3). Coalescent times obtained from BEAST, and ARLEQUIN'S time since expansion, are presented in Figure 6c with details in Table S3. Finally, extended Bayesian skyline plots (Figure 7) are generally indicative of slight population growth in the upper Pleistocene to Holocene.

Nuclear demographic analyses

Nuclear demographic summary statistics generally show mixed results. The Wilcoxon test in BOTTLENECK (Table S3b) was only significant ($P < 0.05$) in the Atlantic under the IAM mutational model, all other tests were non-significant. The M-ratio analyses indicate that the observed M value was smaller than the simulated critical value for both populations indicating that they have undergone recent declines. The only exception was that in the Indo-Pacific the observed value was larger than the critical value when $\Theta = 1$.

The MSVAR analyses (Table S3b) indicate substantial population declines in both of the populations with the point estimate of current effective population size (N_0) almost 2 orders of magnitude less than historic levels (N_1). The declines likely occurred in the Pliocene although the 95% highest posterior densities were quite broad. The average estimated mutation rate across both populations was 3.25×10^{-6}. MCMC diagnostics indicate that the chains were well mixed (PRSF < 1.3) and effective sizes were large (ESS > 200 for each parameter in each population).

Growth rates estimated in LAMARC (Table S3c) were small or negative, and the confidence intervals were close to or spanned 0 indicating a population decline (Kuhner

2006). Estimates of theta were larger than those from the mtCR dataset. Migration rates indicated significantly more gene flow from the Atlantic to the Indo-Pacific. The extended Bayesian skyline plots (Figure 7) indicate a similar decline for both populations in the Holocene. However, although these analyses were run for the maximum number of generations allowed in BEAST, the ESS values were less than 200 so these results should be interpreted cautiously.

Individual assignment analyses

GENECLASS identified no first generation migrants. The STRUCTURE analyses identified one Atlantic individual as having an Atlantic membership coefficient less than 0.85 ($q < 0.78$) although it could not be unambiguously assigned to a class of migrants (posterior probability of being a first, second or third generation migrant of 0.07, 0.23 and 0.11, respectively). All other individuals had membership coefficients > 0.80 in the oceanic basin in which they were sampled. Finally, the single market fin had a mtCR haplotype that clustered with the Atlantic clade in the TCS network (Figure 2), and its microsatellite genotype was assigned to the Atlantic population by both GENECLASS (posterior probability of Atlantic and IndoPacific origin of 0.34 and 0.06, respectively) and STRUCTURE ($q = 0.86$).

Discussion

Contrasting patterns of matrilineal and nuclear population structure

Genetic subdivision of populations allows the identification of management units or genetically distinct lineages (Dizon *et al.* 1992; Waples 1995). Migration rates as low as one individual per generation may be sufficient to obscure differentiation due to genetic drift, although in natural marine populations several immigrants per generation may be necessary to negate this effect (Waples 1998). Additionally, larger numbers of immigrants are necessary to rebuild populations after natural or anthropogenic disturbance (Waples 1998). Statistically significant matrilineal genetic structure was present among eight genetically distinct populations. Thus all eight populations should be considered distinct management units for conservation and fisheries management purposes. Targeted genetic studies with additional animals of known gender and age

class are necessary to define population structure at a finer scale, particularly to fully delineate genetic structure between the eastern and western Atlantic.

Divergence between these maternal lineages occurred very recently, in the Holocene or Upper Pleistocene, with no clear difference between intra- and inter-basin divergence times or population ages across coalescent and mismatch analyses. Female-mediated gene flow between these some of these populations appears to be ongoing but at levels too low to be detected by conventional tagging studies or to counteract population differentiation. Interestingly, these data indicate uni-directional maternal gene flow from the western Atlantic to the eastern Pacific, which is an unusual dispersal route that is not typically reported for sharks.

In contrast, the microsatellite data reveal no significant intra-basin structure with high levels of gene flow among collection areas within oceanic basins. Although there was no statistically significant genetic differentiation between the Atlantic and Indo-Pacific, there was evidence of gene flow between basins with a directional bias again from the Atlantic to the Indo-Pacific.

I corrected for high levels of heterozygosity in microsatellite loci that can depress levels of F_{ST} (Hedrick 1999) by using Mierman's F'_{ST} (Meirmans 2006) which calculates the upper limit for F_{ST} and then scales the observed F_{ST} by the calculated maximum. I also corrected for upward bias in mtCR Φ_{ST}, which accounts for divergence between haplotypes, by calculating the mitochondrial analogue $mtF_{ST,\ which}$ is based only on haplotypic frequency. For diverging populations that otherwise meet the assumptions of Hardy-Weinberg equilibrium, simulations by Larsson *et al.* (2009) show that the ratio of mtDNA F_{ST} to nuclear F_{ST} (R) ranges between 1.0 and 4.0, and that the markers will reach equilibrium in approximately 1000 generations after population separation. At equilibrium, both markers should have equivalent F_{ST} values, although migration can cause the final equilibrium R to rise to close to 4. However, at any point during the divergence process, with or without migration, R is expected to range from 1 – 4, and values less than 1 or greater than 4 indicate significant departure from equilibrium conditions.

I calculated R based on comparisons of both Φ_{ST} and mtF_{ST} with both F_{ST}, and F'_{ST}. Regardless of which mtCR and nuclear statistics were compared, I found strong

indications of non-equilibrium conditions (R >> 10) within oceanic basins. However, the signal is less clear in comparisons between the Atlantic and Indo-Pacific, ranging from $2.7 < R < 6$ for the standardized mtF_{ST} / F'_{ST} comparison to $10 < R < 25$ for the non-standardized Φ_{ST} / F_{ST}.

What could explain the unexpectedly large contrasts in patterns of mitochondrial and nuclear differentiation? Several scenarios can cause deviation from equilibrium conditions, potentially skewing the R ratios, including low effective population size for the mtCR, selection, unequal sex ratios, or male-biased dispersal. The estimates of effective population size are generally comparable between the markers, there was no evidence of selection in calculated values of Fu's F_s and Tajima's D statistics, and reported embryonic sex ratios are 1:1 (Stevens 1984; Castro & Mejuto 1995) although data on adult sex ratios are lacking. This leaves male-mediated gene flow as the most likely explanation for the observed contrast between mitochondrial and nuclear patterns of differentiation, which is consistent with my other results. No intra-basin differentiation was observed with nuclear microsatellites despite significant mtCR structuring, although low levels of intra-basin mtCR gene flow are evident in the haplotypes that are shared across the Atlantic and the Indo-Pacific. The lower inter-basin R ratios also make sense if the dispersal rates between the Atlantic and Indo-Pacific are sufficiently low, and may either retard male dispersal or increase female dispersal by chance or due to an undetermined ecological or selective factor of this dispersal route. Finally, female philopatry to specific or broad nursery areas has been suggested for some shark species, including the great white shark (Pardini *et al.* 2001; Blower *et al.* 2012), the scalloped hammerhead shark (Duncan *et al.* 2006; Nance *et al.* 2011; Daly-Engel *et al.* 2012b), the black tip shark (Keeney *et al.* 2005; Sodré *et al.* 2012), and the blacktip reef shark (Mourier & Planes 2013). Thus this is not the first report of genetic evidence of female philopatry and male mediated gene flow.

Another possible explanation for unexpected patterns of genetic structure is high levels of kinship within a site (Iacchei *et al.* 2013). I found pairs of probably related individuals at all collection areas. The limited size data available indicated that these individuals were probably sub-adults sampled before dispersal from nursery grounds. Although I removed one member of each pair from my analyses, I was only able to test

for parent-offspring, full sibling and half sibling relationships. Thus it is possible that more distantly related individuals are included in the data set, although the effect on F-statistics would likely be minimal. Another alternative hypothesis also deserves mention. Population structure may be influenced by juvenile use of nursery or natal areas if sampling occurs before dispersal (Daly-Engel *et al.* 2012b; Messier *et al.* 2012). There are reported nursery areas for *S. zygaena*, mostly in shallow coastal waters, and many of my sampling areas are near a reported nursery. My samples were obtained opportunistically, with the majority not having gender, size, or specific catch location data. Targeted collection of individuals of known size class and gender would be necessary to definitively test these two alternatives. However, male-mediated gene flow is the more likely explanation because mtCR structure is indicated when only definite adult size classes are included (data not shown), although the samples sizes are too small to be statistically significant.

In comparison, studies of a sister species, *S. lewini*, revealed similar patterns. *S. lewini* exhibits comparable mtCR population structure on a global scale (Duncan *et al.* 2006), although more fine-scale structure is present at least in certain areas (Nance *et al.* 2011). Nuclear divergence between the Atlantic and Indo-Pacific is also significant with slight intra-basin structure present particularly across large oceanic expanses (Daly-Engel *et al.* 2012b). A similar pattern was also found in another sister species, *Sphyrna mokarran* (Testerman, Chapter 3), with strong geographic partitioning between the Atlantic and Australia and little intra-basin structure with both mtCR and microsatellites, although there were fine scale differences between the two marker types involving samples from the northern Indian Ocean. Variability in genetic signatures can occur even among closely related species (Rocha *et al.* 2002; Gaither *et al.* 2010; DiBattista *et al.* 2013) and may be related to innate differences in life history or ecological preferences. It is possible that the differences in the patterns of genetic variation in these three large bodied hammerheads are most likely due to varying levels of female philopatry and male mediated gene flow. Continental barriers between oceanic basins present variable permeability for the three species, depending on thermal tolerances and past oceanic connectivity (Testerman, Chapter 3).

Shallow population structure and unusual Atlantic / Indo-Pacific connection

Fossil evidence indicates a species origin in the Miocene (Cappetta 1987). This is consistent with the molecular phylogeny of Lim *et al.* (2010) that suggests that extant *Sphyrna* species have diversified within the past 10 million years. My mtCR data indicate coalescent and divergence times that are much more recent. Assuming a 20-year generation time, these estimated coalescent and divergence times are generally more than the approximately 1 000 generations necessary to achieve equilibrium after population separation. The relatively shallow population structure evident in the TCS haplotype network is consistent with recent divergences. The contrast between my data and fossil and molecular phylogenies could be due to several factors, including low effective population sizes, severe population bottlenecks, continuous gene flow, or recent rapid dispersal and subsequent isolation. Except for the western North Atlantic, my estimated population sizes are generally large. I suspect that a combination of historic bottlenecks and recent dispersal with low levels of ongoing gene flow have contributed to the shallow population structure observed here. The contrasting signals of population expansion based on mtCR sequences and decline based on microsatellite data are not unusual, and lend support to a combination of demographic events contributing to shallow population structure. Similar contrasting patterns have been reported in species as diverse as the scalloped hammerhead (Nance *et al.* 2011) and the giant panda (Hu *et al.* 2010). Indeed, these results are consistent with the expectation that statistical tests and coalescent analyses based on microsatellite data are more sensitive to recent events while those based on mtCR detect more ancient events. Additionally, my calculated microsatellite mutation rate seems reasonable given that reported microsatellite mutation rates for fish are in the range of 10^{-4} to 10^{-5} (Shimoda *et al.* 1999; Yue *et al.* 2007), the reported range for mammals is 10^{-2} to 10^{-5} (Ellegren 2004), and mitochondrial and nuclear DNA appears to mutate approximately an order of magnitude more slowly in sharks than in other taxa (Martin *et al.* 1992; Martin 1999).

Despite this observed shallow population structure, I see hints of historical events in the species genetic signature. In the mtCR haplotype network, the ancestral haplotype is observed primarily in the western Indian Ocean. The eastern North Pacific clade is connected but separate from the other clade and the Atlantic clade is terminal. This

arrangement is suggestive of unidirectional dispersal from the western Indian Ocean to the western Pacific, the eastern Pacific, and then the Atlantic around the southern tip of South America, although my estimated migration rates suggest dispersal from the western Atlantic to the eastern South Pacific. However, my coalescent and mismatch estimates of population age and divergence times are oldest and similar for the western Indian Ocean, eastern North Pacific and western Atlantic, lending support to my hypothesis of recent bottlenecks, ongoing migration and/or recent divergence obscuring more ancient historical population structure.

Other studies have suggested connectivity between the Atlantic and Indo-Pacific via South America in antitropical taxa (Schwaninger 2008; Verissimo *et al.* 2010; Herrera *et al.* 2012), although this is an atypical dispersal route for most sharks. This inferred migratory pattern could result from passive transport down the eastern coast of South America and back up into the eastern South Pacific by the coastal counter currents associated with the South Atlantic Current. The South Atlantic Current itself could also be an offshore mechanism of passive transport from the western South Pacific to the eastern South Atlantic. The thermal tolerance of *S. zygaena* includes sea surface temperatures observed around the southern edge of South America (Guhin *et al.* 2003) and the species geographic range includes this region (Last & Stevens 2009), so it is not really surprising to find genetic evidence of this dispersal pathway. Further study that incorporates additional samples from the eastern South Atlantic is necessary to assess the relative importance of this dispersal route versus potential connectivity via southern Africa.

Conclusions

This is the first global phylogeographic study of *S. zygaena*. Three broad conclusions are apparent from the data. First, there is significant matrilineal genetic structure in all pairwise comparisons, with no intra-basin nuclear genetic differentiation. These patterns of genetic variation should enable genetic tests of geographic origin similar to those described in previous chapters. Secondly, the contrasting phylogeographic patterns observed between mtCR and nuclear microsatellite data indicate the probable occurrence of female philopatry and male mediated gene flow. Finally, there is genetic evidence of a

dispersal pathway between the eastern South Pacific and western Atlantic that although unusual for sharks, is not unexpected for this amphitemperate species.

Acknowledgements

This work could not have been conducted without the collection of samples of *S. zygaena*. For their help in sample collection I thank D. Abercrombie, J. Bizzaro, D. Cartamil, J. Cassin, J. Castro, D. Chapman, G. Cliff, C. Conrath, D. Grubbs, S. Gulak, F. Hazin, S. Hernandez, E. Jones, M. Joung, D. Kacev, T. Leary, M. Liao, S. Makien, P. Mancini, C. Martins, R. McCauley, P. Motta, H. Nance, L. Natanson, L. Noll, D. Pinhal, J. Quattro, D. Reid, W. Robbins, M. Sampson, and S. Wintner. I also thank J. Hester for her assistance with laboratory work. Financial support for this project was received from Guy Harvey Ocean Foundation, Save Our Seas Foundation and SeaGrant Florida.

References

Alter SE, Ramirez SF, Nigenda S, Ramirez JU, Bracho LR, Palumbi SR (2009) Mitochondrial and nuclear genetic variation across calving lagoons in eastern North Pacific gray whales (*Eschrichtius robustus*). Journal of heredity, 100, 34-46.

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Moleclar Biology and Evolution, 13, 494-504.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire génome, populations, interactions, CNRS UMR, 5000, 1996-2004.

Blower D, Pandolfi J, Bruce B, Gomez-Cabrera M, Ovenden J (2012) Population genetics of Australian white sharks reveals fine-scale spatial structure, transoceanic dispersal events and low effective population sizes. Marine Ecology Progress Series, 455, 229-244.

Botello A, Iliffe TM, Alvarez F, Juan C, Pons J, Jaume D (2013) Historical biogeography and phylogeny of *Typhlatya* cave shrimps (Decapoda: Atyidae) based on mitochondrial and nuclear data. Journal of Biogeography, 40, 594-607.

Bowen BW, Bass AL, Soares L, Toonen RJ (2005) Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). Molecular Ecology.

Bowen BW, Karl SA (2007) Population genetics and phylogeography of sea turtles. Molecular Ecology, 16, 4886-4907.

Cappetta H (1987) Chondrichthyes II: mesozoic and cenozoic elasmobranchii Gustav Fischer Verlag, Stuttgart, Germany.

Casado-Amezua P, Goffredo S, Templado J, Machordom A (2012) Genetic assessment of population structure and connectivity in the threatened Mediterranean coral *Astroides calycularis* (Scleractinia, Dendrophylliidae) at different spatial scales. Molecular Ecology, 21, 3671-3685.

Casper BM*, et al.* (2005) *Sphyrna zygaena*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. IUCN Red List.

Castro ALF, Stewart BS, Wilson SG, Hueter RE, Meekan MG, Motta PJ, Bowen BW, Karl SA (2007) Population genetic structure of Earth's largest fish, the whale shark (Rhincodon typus). Molecular Ecology, 16, 5183-5192.

Castro J, Mejuto J (1995) Reproductive parameters of blue shark, *Prionace glauca*, and other sharks in the Gulf of Guinea. Marine and Freshwater Research, 46, 967-973.

Chapman DD, Babcock EA, Gruber SH, DiBattista JD, Franks BR, Kessel SA, Guttridge T, Pikitch EK, Feldheim KA (2009) Long-term natal site-fidelity by immature lemon sharks (*Negaprion brevirostris*) at a subtropical island. Molecular Ecology, 18, 3500-3507.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9, 1657-1659.

Coelho R, Fernandez-Carvalho J, Lino PG, Santos MN (2012) An overview of the hooking mortality of elasmobranchs caught in a swordfish pelagic longline fishery in the Atlantic Ocean. Aquatic Living Resources, 25, 311-319.

Compagno LJV (1984) FAO Species Catalog. Vol. 4 Sharks of the world, Part 2 - Carcharhiniformes. In: FAO Species Catalog. Food and Agriculture Organization of the United Nations, Rome.

Coombs JA, Letcher BH, Nislow KH (2008) CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. Molecular Ecology Resources, 8, 578-580.

Cortés E*, et al.* (2010) Ecological risk assessment of pelagic sharks caught in Atlantic pelagic longline fisheries. Aquatic Living Resources, 23, 25-34.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics, 134, 959-969.

Daly-Engel T, Randall J, Bowen B (2012a) Is the Great Barracuda (*Sphyraena barracuda*) a reef fish or a pelagic fish? The phylogeographic perspective. Marine Biology, 159, 975-985.

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW (2012b) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). PLOS One, 7, e29986.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 1-38.

Díaz-Jaimes P, Uribe-Alcocer M, Rocha-Olivares A, García-de-León FJ, Nortmoon P, Durand JD (2010) Global phylogeography of the dolphinfish (*Coryphaena hippurus*): The influence of large effective population size and recent dispersal on the divergence of a marine pelagic cosmopolitan species. Molecular Phylogenetics and Evolution, 57, 1209-1218.

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, Choat JH, Craig MT, Skillings DJ, Bowen BW (2013) After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. Journal of Biogeography, 40, 1170-1181.

Diemer KM, Mann BQ, Hussey NE (2011) Distribution and movement of scalloped hammerhead *Sphryna lewini* and smooth hammerhead *Sphyrna zygaena* sharks along the east coast of southern Africa. African Journal of Marine Science, 33, 229-238.

Dizon AE, Lockyer C, Perrin WF, DeMaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. Conservation Biology, 6, 24-36.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7, 214.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2008) Geneious v4.0.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29, 1969-1973.

Duchesne P, Turgeon J (2009) FLOCK: a method for quick mapping of admixture without source samples. Molecular Ecology Resources, 9, 1333-1344.

Duchesne P, Turgeon J (2012) FLOCK Provides Reliable Solutions to the "Number of Populations" Problem. Journal of heredity, 103, 734-743.

Dudgeon CL, Broderick D, Ovenden JR (2009) IUCN classification zones concord with, but underestimate, the population genetic structure of the zebra shark *Stegostoma fasciatum* in the Indo-West Pacific. Molecular Ecology, 18, 248-261.

Duncan K, Martin A, Bowen B, DeCouet H (2006) Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*). Molecular Ecology, 15, 2239-2251.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Molecular Ecology, 11, 2571-2581.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359-361.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. Theoretical applications in genetics, 92, 832-839.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics, 5, 435-445.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics, 2005, 47-50.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131, 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics, 164, 1567-1587.

Fredsted PV (2006) FaBox - an online fasta sequence toolbox, http://www.birc.au.dk/fabox.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147, 919-925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics, 133, 693-709.

Gaida IH (1997) Population Structure of the Pacific Angel Shark, *Squatina californica* (Squatiniformes: Squatinidae), around the California Channel Islands. Copeia, 1997, 738-744.

Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). Journal of Biogeography, 37, 133-147.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology, 10, 305-318.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of heredity, 86, 485-486.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Guhin S, Ray P, Mariano AJ, Ryan E (2003) The South Atlantic Current. In: Ocean Surface Currents.

Guo SW, Thompson EA (1992) Performing exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics, 48, 361-372.

Harpending R (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology, 66, 591-600.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. Evolution, 53, 313-318.

Hedrick PW (2005) A standardized genetic differentiation measure Evolution, 59, 1633-1638.

Herrera S, Shank TM, Sánchez JA (2012) Spatial and temporal patterns of genetic variation in the widespread antitropical deep-sea coral *Paragorgia arborea*. Molecular Ecology, 21, 6053-6067.

Hoelzel AR, Shivji MS, Magnussen JE, Francis MP (2006) Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*). Biology Letters, 2, 639-642.

Hu Y, Qi D, Wang H, Wei F (2010) Genetic evidence of recent population contraction in the southernmost population of giant pandas. Genetica, 138, 1297-1306.

Hueter R, Heupel M, Heist E, Keeney D (2005) Evidence of philopatry in sharks and implications for the management of shark fisheries. e-Journal of Northwest Atlantic Fishery Science, 35, 239-247.

Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, García-Rodríguez FJ, Toonen RJ (2013) Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Molecular Ecology, 22, 3476-3494.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, 23, 1801-1806.

Kalinowski ST, Wagner AP, Taper ML (2006) ml-relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes, 6, 576-579.

Keeney D, Heupel M, Hueter R, Heist E (2005) Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. Molecular Ecology, 14, 1911-1923.

Kohler N, Turner P (2001) Shark tagging: a review of conventional methods and studies. Environmental Biology of Fishes, 60, 191-223.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics, 22, 768-770.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular genetics analysis and sequence alignment. Briefings in Bioinformatics, 5, 150-163.

Larsson L, Charlier J, Laikre L, Ryman N (2009) Statistical power for detecting genetic divergence—organelle versus nuclear markers. Conservation Genetics, 10, 1255-1264.

Last PR, Stevens JD (2009) Sharks and rays of Australia, 2nd edn. Harvard University Press, Cambridge, Massachusetts.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451-1452.

Lim DD, Motta P, Mara K, Martin AP (2010) Phylogeny of hammerhead sharks (Family Sphyrnidae) inferred from mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution, 55, 572-579.

Martin A, Naylor G, Palumbi S (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature, 357, 1992.

Martin AP (1999) Substitution rates of organelle and nuclear genes in sharks: implicating metabolic rate (again). Molecular Biology and Evolution, 16, 996-1002.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution, 60, 2399-2402.

Mendonça F, Oliveira C, Gadig OF, Foresti F (2011) Phylogeography and genetic population structure of Caribbean sharpnose shark *Rhizoprionodon porosus*. Reviews in Fish Biology and Fisheries, 21, 799-814.

Messier GD, Garant D, Bergeron P, Réale D (2012) Environmental conditions affect spatial genetic structures and dispersal patterns in a solitary rodent. Molecular Ecology, 21, 5363-5373.

Mourier J, Planes S (2013) Direct genetic evidence for reproductive philopatry and associated fine-scale migrations in female blacktip reef sharks (*Carcharhinus melanopterus*) in French Polynesia. Molecular Ecology, 22, 201-214.

Nance HA, Daly-Engel TS, Marko PB (2009) New microsatellite loci for the endangered scalloped hammerhead shark, *Sphyrna lewini*. Molecular Ecology Resources, doi: 10.1111/j.1755-0998.2008.02510.x.

Nance HA, Klimley P, Galvan-Magana F, Martinez-Ortiz J, Marko PB (2011) Demographic Processes Underlying Subtle Patterns of Population Structure in the Scalloped Hammerhead Shark, *Sphyrna lewini*. PLOS One, 6, e21459.

Narum SR, Stephenson JJ, Campbell MR (2007) Genetic Variation and Structure of Chinook Salmon Life History Types in the Snake River. Transactions of the American Fisheries Society, 136, 1252-1262.

Nei M (1987) Molecular evolutionary genetics Columbia University Press, New York.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a markov chain monte carlo approach (MDIV). Genetics, 158, 885-896.

Ovenden JR, Dashiwagi T, Broderick D, Giles J, Salini J (2009) The extent of population genetic subdivision differs among four co-distributed shark species in the Indo-Australian archipelago. BMC Evolutionary Biology, 9, 40.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology, 13, 55-65.

Pardini A*, et al.* (2001) Sex-biased dispersal of great white sharks. Nature, 412, 139-140.

Park S (2001) Trypanotolerance in West African cattle and the population genetic effects of selection, PhD Dissertation, University of Dublin, Dublin, Ireland.

Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics, 28, 2537-2539.

Peakall ROD, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288-295.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration and secondary contact. Evolution, 56, 1776-1788.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of heredity, 95, 536-539.

Piry S, Luikart G, Cornuet J-M (1999) BOTILENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of heredity, 90, 502-503.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics, 14, 817-818.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA, 94, 9197-9201.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol. Biol. Evol., 20, 76-86.

Rice WR (1989) Analyzing tables of statistical tests. Evolution, 43, 223-225.

Richards V, Henning M, Witzell W, Shivji M (2009) Species delineation and evolutionary history of the globally distributed spotted eagle ray (*Aetobatus narinari*). Journal of heredity, 100, 273-283.

Rocha LA, Bass AL, Robertson R, Bowen BW (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). Molecular Ecology, 11, 243-252.

Rocha LA, Craig MT, Bowen BW (2007) Phylogeography and the conservation of coral reef fishes. Coral Reefs, 26, 513-513.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution, 9, 552-569.

Rosenberg NA (2004) distruct: a program for the graphical display of population structure. Molecular Ecology Notes, 4, 137-138.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8, 103-106.

Rousset F, Raymond M (1995) Testing heterozygote excess and deficiency. Genetics, 140, 1413-1419.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. Genetics, 152, 1079-1089.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18, 233-234.

Schwaninger HR (2008) Global mitochondrial DNA phylogeography and biogeographic history of the antitropically and longitudinally disjunct marine bryozoan *Membranipora membranacea* L. (Cheilostomata): Another cryptic marine sibling species complex? Molecular Phylogenetics and Evolution, 49, 893-908.

Shimoda N*, et al.* (1999) Zebrafish Genetic Map with 2000 Microsatellite Markers. Genomics, 58, 219-232.

Smale MJ, Cliff G (1998) Cephalopods in the diets of four shark species (*Galeocerdo cuvier, Sphyrna lewini, S. zygaena* and *S. mokarran)* from KwaZulu-Natal, South Africa. South African Journal of Marine Science, 20, 241-253.

Sodré D, Rodrigues-Filho LF, Souza RF, Rêgo PS, Schneider H, Sampaio I, Vallinoto M (2012) Inclusion of South American samples reveals new population structuring of the blacktip shark (*Carcharhinus limbatus*) in the western Atlantic. Genetics and Molecular Biology, 35, 752-760.

Stevens J (1984) Biological observations on sharks caught by sport fisherman of New South Wales. Marine and Freshwater Research, 35, 573-590.

Storz JF, Beaumont MA (2002) Testing for genetic evidence of population expansion and contraction: An empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. Evolution, 56, 154-166.

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. Genetics, 143, 1457-1465.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Moleclar Biology and Evolution, 10, 512-526.

Templeton A, Crandall K, Sing C (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics, 132, 619-633.

Testerman CB, Fitzpatrick S, Prodohl PA, Shivji M (In preparation) Development and characterization of novel microsatellite loci for the great hammerhead shark *Sphyrna mokarran* and their cross-species amplification among other sphyrnid species.

Theisen TC, Bowen BW, Lanier W, Baldwin JD (2008) High connectivity on a global scale in the pelagic wahoo, *Acanthocybium solandri* (tuna family Scombridae). Molecular Ecology, 17, 4233-4247.

Tillett BJ, Meekan MG, Field IC, Thorburn DC, Ovenden JR (2012) Evidence for reproductive philopatry in the bull shark *Carcharhinus leucas*. Journal of Fish Biology, 80, 2140-2158.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535-538.

Verissimo A, McDowell JR, Graves JE (2010) Global population structure of the spiny dogfish *Squalus acanthias*, a temperate shark with an antitropical distribution. Molecular Ecology, 19, 1651-1662.

Waples RS (1995) Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act 17, 8-27.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. The Journal of Heredity, 89, 438-450.

Whitney NM, Robbins WD, Schultz JK, Bowen BW, Holland KN (2012) Oceanic dispersal in a sedentary reef shark (*Triaenodon obesus*): genetic evidence for extensive connectivity without a pelagic larval stage. Journal of Biogeography, 39, 1144-1156.

Wiley TR, Simpfendorfer CA (2007) The ecology of elasmobranchs occurring in the Everglades National Park, Florida: implications for conservation and management. Bulletin of Marine Science, 80, 171-189.

Wright S (1951) The genetical structure of populations. Annals of Eugenics, 15, 323-354.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, 19, 395-420.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). Genetica, 129, 329-331.

Tables

Table 1. Population statistics for *Sphyrna zygaena*. Number of individuals (*n*), number of haplotypes (*nh*), haplotype diversity (*h*), standard deviation (*SD*), nucleotide diversity (π). Collection areas are coded as follows: (i) Atlantic (ATL): western North Atlantic (WNA), western South Atlantic (WSA); (ii) Indo Pacific (IP): western Indian Ocean (WIO), western South Pacific (WSP), western North Pacific (WNP), eastern North Pacific (ENP), eastern Tropical Pacific (ETP), and eastern South Pacific (ESP).

	Atlantic (ATL)		Pooled	Indo Pacific (IP)						Pooled	Pooled All
	WNA	WSA	ATL	WIO	WSP	WNP	ENP	ETP	ESP	IP	Samples
n	21	55	87	63	44	11	55	15	26	216	303
nh	2	4	5	15	4	4	4	4	7	26	31
h	0.1810	0.4357	0.3692	0.6933	0.4038	0.6909	0.3434	0.6381	0.6800	0.8733	0.8841
SD	0.1044	0.0695	0.0599	0.0613	0.0770	0.1276	0.0746	0.0931	0.0858	0.0098	0.0085
π	0.0000	0.0007	0.0005	0.0016	0.0007	0.0010	0.0008	0.0007	0.0011	0.0025	0.0032
SD	0.0000	0.0006	0.0005	0.0010	0.0006	0.7743	0.0006	0.0006	0.0008	0.0015	0.0018

Table 2. AMOVA, SAMOVA, and pairwise Φ_{ST}. The amount of mtCR genetic diversity within and among groups.

Table 2a. Pairwise Φ_{ST}, all are significant at $P < 0.00001$.

	WNA	WSA	WIO	WSP	WNP	ENP	ETP	ESP
Western North Atlantic (WNA)								
Western South Atlantic (WSA)	0.1116							
Western Indian Ocean (WIO)	0.7857	0.7697						
Western South Pacific (WSP)	0.8852	0.8469	0.5290					
Western North Pacific (WNP)	0.8761	0.7667	0.4864	0.5351				
Eastern North Pacific (ENP)	0.9067	0.8727	0.3581	0.7702	0.7625			
Eastern Tropical Pacific (ETP)	0.8735	0.7553	0.6524	0.7462	0.4632	0.8345		
Eastern South Pacific (ESP)	0.7596	0.6991	0.5838	0.6778	0.2912	0.7856	0.0949	

Table 2b. Heirarchical AMOVA.

		Φ Statistic	
	% variation	Value	P value
Among groups $\mathbf{\Phi}_{CT}$	49.10%	0.4991	0.0430
Among populations within groups Φ_{SC}	31.68%	0.6325	0.0000
Within populations Φ_{ST}	18.41%	0.8159	0.0000

Table 2c. SAMOVA, K=8.

		Φ Statistic	
	% variation	Value	P value
Among groups $\mathbf{\Phi}_{CT}$	78.14%	0.7814	0.0205
Among populations within groups Φ_{SC}	-0.85%	-0.0388	0.0000
Within populations Φ_{ST}	22.71%	0.7729	0.0000

Table 3. Net pairwise genetic distances.

Table 3a. Among *S. zygaena* populations.

	WNA	WSA	WIO	WSP	WNP	ENP	ETP
Western North Atlantic (WNA)							
Western South Atlantic (WSA)	0.00004						
Western Indian Ocean (WIO)	0.0038	0.0036					
Western South Pacific (WSP)	0.0038	0.0039	0.0008				
Western North Pacific (WNP)	0.0023	0.0024	0.0009	0.0009			
Eastern North Pacific (ENP)	0.0050	0.0047	0.0007	0.0019	0.0020		
Eastern Tropical Pacific (ETP)	0.0021	0.0021	0.0021	0.0021	0.0007	0.0033	
Eastern South Pacific (ESP)	0.0019	0.0019	0.0017	0.0018	0.0004	0.0029	0.0001

Table 3b. Among *Sphyrna* species.

	Eblo	*Slew*	*Smok*
E. blochii (Eblo)			
S. lewini (Slew)	0.1654		
S. mokarran (Smok)	0.1712	0.1967	
S. zygaena (Szyg)	0.1382	0.1421	0.1449

Table 4. *S. zygaena* summary statistics for each microsatellite locus by collection area. n = number of individuals, Na = number of alleles, A_r = allelic richness, GD = genetic diversity (Nei, 1987), H_E = expected frequency of heterozygotes, H_O = observed frequency of heterozygotes, Null = frequency of null alleles estimated using the Expectation Maximization (EM) algorithm of Dempster, Laird, and Rubin (1977). Far right column contains for each locus across all populations: (i) mean H_E, H_O values, (ii) overall A_r and F_{IS}, and (iii) total Na and null allele frequency. The bottom row contains for each population across all loci: (i) mean number of alleles, (ii) observed and expected frequency of heterozygotes, (iii) F_{IS}, (iiii) mean null allele frequency. Last two lines of the bottom row, far right column are the overall F_{IS} and mean null allele frequency for all loci in all populations. Bold F_{IS} values are significant after sequential Bonferroni correction ($P < 0.05$), bold null allele values are > 0.10. Collection localities are coded as in Table 1.

Table 4. (Continued).

Locus	WNA *n* = 21	WSA *n* = 53	WIO *n* = 103	WSP *n* = 45	WNP *n* = 11	ENP *n* = 52	ETP *n* = 15	ESP *n* = 17	Mean/locus Total/locus
GhhA62									
Na	4	4	6	4	4	8	4	4	8
A_r	3.3310	2.8010	2.9850	2.8670	3.7900	3.6860	3.7940	3.3760	3.2500
GD	0.5120	0.2580	0.3090	0.4620	0.4590	0.4060	0.5930	0.4560	0.4370
H_e	0.5000	0.2558	0.3071	0.4553	0.4421	0.4015	0.5711	0.4434	0.3825
H_o	0.5238	0.2453	0.2913	0.3636	0.5455	0.3750	0.5333	0.5000	0.3558
F_{IS}	-0.0230	0.0510	0.0570	0.2120	-0.1880	0.0760	0.1000	-0.0960	0.0787
Null	0.0523	0.0881	0.0235	0.0846	0.0000	0.0408	0.0000	0.0000	0.0362
GhhA8									
Na	10	20	28	29	6	20	12	12	56
A_r	6.8830	7.5220	8.3280	9.7280	5.7950	7.7240	8.4970	8.6640	8.1780
GD	0.7990	0.8020	0.8320	0.8630	0.7560	0.8340	0.8330	0.8020	0.8200
H_e	0.7823	0.7941	0.8276	0.8533	0.7200	0.8247	0.8022	0.7679	0.8271
H_o	0.9048	0.7647	0.8119	0.8667	0.8000	0.7451	0.7333	0.6429	0.7864
F_{IS}	-0.1330	0.0470	0.0240	-0.0040	-0.0590	0.1060	0.1200	0.1990	0.0508
Null	0.0000	0.0000	0.0062	0.0000	0.0000	0.0358	0.0333	0.0125	0.0110
GhhC7									
Na	6	7	9	6	6	7	7	4	12
A_r	4.8370	4.3270	4.3680	4.1580	5.4540	3.9730	5.6860	3.8070	4.5550
GD	0.7700	0.7000	0.6820	0.7020	0.7360	0.6200	0.6860	0.6880	0.7010
H_e	0.7506	0.6942	0.6782	0.6925	0.7025	0.6129	0.6644	0.6660	0.6914
H_o	0.7143	0.8077	0.6667	0.6250	0.7273	0.6304	0.7333	0.6875	0.6901
F_{IS}	0.0730	-0.1540	0.0220	0.1100	0.0120	-0.0180	-0.0690	0.0000	0.0035
Null	0.0077	0.0000	0.0000	0.0298	0.0000	0.0041	0.0000	0.0000	0.0052
CL12									
Na	14	18	30	23	15	25	12	16	44
A_r	8.5960	8.1760	10.6690	11.1370	12.9660	11.2850	9.7280	11.3670	10.5700
GD	0.8540	0.8400	0.9190	0.9350	0.9550	0.9310	0.9020	0.9390	0.9200
H_e	0.8345	0.8327	0.9141	0.9235	0.9132	0.9225	0.8711	0.9118	0.9149
H_o	0.9048	0.8679	0.8725	0.8667	1.0000	0.9615	0.8667	0.9412	0.8912
F_{IS}	-0.0600	-0.0330	0.0500	0.0730	-0.0480	-0.0330	0.0400	-0.0020	0.0273
Null	0.0000	0.0000	0.0228	0.0435	0.0000	0.0000	0.0000	0.0000	0.0083
CL100									
Na	6	9	10	9	4	9	7	9	12
A_r	4.8810	5.5570	5.9440	6.2290	3.9970	5.8990	5.7800	7.2620	5.8530
GD	0.7720	0.7680	0.8000	0.8140	0.7360	0.8000	0.7980	0.8250	0.8010
H_e	0.7525	0.7605	0.7960	0.8039	0.7066	0.7927	0.7733	0.8045	0.802
H_o	0.7500	0.7451	0.7379	0.8095	0.8182	0.8654	0.8667	0.9412	0.7853
F_{IS}	0.0290	0.0300	0.0780	0.0050	-0.1110	-0.0820	-0.0870	-0.1400	0.0223
Null	0.0179	0.0029	0.0367	0.0000	0.0000	0.0000	0.0000	0.0000	0.0072
GhhD1									
Na	28	47	62	45	17	48	26	21	76
A_r	15.3050	15.2330	15.3500	15.4000	15.5840	15.3220	16.2300	14.5330	15.5100
GD	0.9830	0.9810	0.9810	0.9820	0.9830	0.9810	0.9860	0.9740	0.9830
H_e	0.9563	0.9706	0.9758	0.9704	0.9350	0.9706	0.9511	0.9422	0.9804
H_o	0.9000	0.8868	0.8738	0.9111	1.0000	0.9216	0.9333	1.0000	0.9083
F_{IS}	0.0840	**0.0960**	**0.1090**	0.0720	-0.0170	0.0600	0.0530	-0.0270	**0.0752**
Null	0.0263	0.0415	0.0553	0.0309	0.0000	0.0228	0.0000	0.0000	0.0221
GhhD9									
Na	3	2	3	3	2	3	2	2	4
A_r	2.2520	1.9880	2.2340	2.6380	2.0000	2.3920	2.0000	1.9840	2.2810
GD	0.1800	0.3200	0.3810	0.3390	0.3440	0.2970	0.4670	0.2570	0.3260
H_e	0.1757	0.3177	0.3793	0.3360	0.3200	0.2946	0.4444	0.2509	0.337
H_o	0.1905	0.3585	0.3981	0.4000	0.2000	0.3077	0.2667	0.2941	0.3505
F_{IS}	-0.060	-0.119	-0.044	-0.179	0.419	-0.035	0.429	-0.143	-0.0385
Null	0.0000	0.0000	0.0000	0.0000	**0.1106**	0.0000	**0.1291**	0.0000	0.0300
GhhD103									
Na	11	15	14	14	9	14	9	7	20
A_r	8.6360	8.9620	7.5880	7.8360	8.6950	8.3340	7.7980	6.4310	8.5580
GD	0.8930	0.8950	0.8510	0.8530	0.9220	0.8700	0.8740	0.8440	0.8850
H_e	0.8726	0.8873	0.8465	0.8438	0.8700	0.8621	0.8393	0.8114	0.8807
H_o	1.0000	0.9811	0.8333	0.9048	0.8000	0.8654	0.7857	0.5882	0.8642
F_{IS}	-0.1190	-0.0960	0.0200	-0.0600	0.1330	0.0060	0.1010	0.3030	0.0203
Null	0.0000	0.0000	0.0083	0.0000	0.0333	0.0096	0.0354	**0.1127**	0.0249

Table 4. (Continued).

Locus	WNA *n* = 21	WSA *n* = 53	WIO *n* = 103	WSP *n* = 45	WNP *n* = 11	ENP *n* = 52	ETP *n* = 15	ESP *n* = 17	Mean/locus Total/locus
GhhF3									
Na	2	2	4	3	2	2	2	2	4
A_r	1.6790	1.8850	2.1110	2.1920	1.8180	1.9540	1.9940	2.0000	2.0270
GD	0.0930	0.1880	0.2530	0.3550	0.0910	0.2490	0.2860	0.4490	0.2560
H_e	0.0907	0.1860	0.2515	0.3504	0.0868	0.2469	0.2778	0.4377	0.2529
H_o	0.0952	0.2075	0.2913	0.3111	0.0909	0.2885	0.3333	0.5294	0.2711
F_{IS}	-0.0260	-0.1060	-0.1530	0.1230	0.0000	-0.1590	-0.1670	-0.1800	-0.0703
Null	0.0000	0.0000	0.0000	0.0358	0.0000	0.0000	0.0000	0.0000	0.0045
Sle27									
Na	8	13	14	10	5	10	5	6	17
A_r	6.7680	7.1500	4.6170	4.3890	4.2730	3.3320	4.3760	3.6470	5.5690
GD	0.8450	0.8110	0.4650	0.4510	0.3410	0.2670	0.5400	0.2760	0.5850
H_e	0.8231	0.8040	0.4633	0.4459	0.3223	0.2644	0.5222	0.2682	0.5983
H_o	0.7619	0.8491	0.4854	0.4667	0.2727	0.2885	0.5333	0.2941	0.5241
F_{IS}	0.0990	-0.0470	-0.0430	-0.0350	0.2000	-0.0810	0.0130	-0.0670	**0.1255**
Null	0.0000	0.0057	0.0133	0.0000	0.0000	0.0000	0.0330	0.0000	0.0065
Sle28									
Na	4	6	8	5	5	5	3	5	10
A_r	3.3590	3.3680	4.0560	3.4960	4.7000	3.7500	2.6000	3.9410	3.6730
GD	0.5880	0.5410	0.5730	0.4460	0.6170	0.5600	0.5020	0.5880	0.5460
H_e	0.5760	0.5358	0.5696	0.4389	0.5700	0.5510	0.4867	0.5625	0.5480
H_o	0.6667	0.5294	0.5730	0.4545	0.3000	0.6538	0.5333	0.3750	0.5418
F_{IS}	-0.134	0.022	0	-0.02	0.514	-0.168	-0.062	0.362	0.0130
Null	0.0000	0.0378	0.0155	0.0000	0.1609	0.0000	0.0000	**0.1082**	0.0403
Sle71									
Na	3	3	3	4	2	3	3	4	4
A_r	2.9870	2.6060	2.8020	3.1630	2.0000	2.8600	2.9820	3.3130	2.9740
GD	0.6040	0.4130	0.4060	0.4560	0.3640	0.4680	0.5880	0.4720	0.4810
H_e	0.5907	0.4092	0.4042	0.4501	0.3512	0.4635	0.5711	0.4550	0.4539
H_o	0.6667	0.4151	0.4118	0.4000	0.4545	0.5098	0.6667	0.3529	0.4576
F_{IS}	-0.1050	-0.0050	-0.0140	0.1220	-0.2500	-0.0900	-0.1340	0.2530	-0.0066
Null	0.2422	0.2140	0.0904	0.0869	0.0000	0.0861	0.0330	**0.1174**	0.1088
Sle77									
Na	9	9	11	10	6	13	5	9	14
A_r	6.7180	5.9920	7.1980	6.7890	5.6950	6.7990	4.1820	7.1480	6.8680
GD	0.7660	0.7720	0.8350	0.8260	0.6830	0.8090	0.6400	0.8350	0.7860
H_e	0.7450	0.7644	0.8315	0.8178	0.6400	0.8007	0.6244	0.8080	0.8133
H_o	0.7000	0.7885	0.8641	0.9111	0.5000	0.7885	0.8000	0.7647	0.8116
F_{IS}	0.0860	-0.0220	-0.0340	-0.1030	0.2680	0.0250	-0.2490	0.0840	0.0037
Null	0.0000	0.0000	0.0000	0.0000	0.1156	0.0220	0.0000	0.0000	0.0172
Sle86									
Na	3	4	4	6	4	4	3	4	7
A_r	2.7300	2.8890	3.6220	4.1660	3.8160	3.4160	3.0000	3.8990	3.6760
GD	0.5120	0.4830	0.6470	0.6860	0.6770	0.6190	0.6740	0.6880	0.6310
H_e	0.4986	0.4785	0.6435	0.6774	0.6405	0.6119	0.6578	0.6660	0.6191
H_o	0.5263	0.5294	0.5889	0.6512	0.5455	0.5294	0.8667	0.6875	0.5877
F_{IS}	-0.0290	-0.0970	0.0900	0.0500	0.1950	0.1440	-0.2860	0.0000	0.0524
Null	0.0291	0.0217	0.0367	0.0350	0.0413	0.0425	0.0296	0.0325	0.0335
Tgr47									
Na	5	7	8	7	5	8	4	6	11
A_r	4.3660	4.3550	4.6180	4.8010	5.0000	5.3570	3.8480	5.1010	4.8120
GD	0.6350	0.5680	0.7430	0.7300	0.7640	0.7800	0.7210	0.7700	0.7440
H_e	0.6191	0.5616	0.7391	0.7216	0.7346	0.7718	0.6956	0.7457	0.7443
H_o	0.7500	0.5510	0.7204	0.7750	1.0000	0.7800	0.6667	0.7059	0.7191
F_{IS}	-0.1800	0.0290	0.0310	-0.0610	-0.3090	-0.0010	0.0760	0.0840	0.0356
Null	0.0000	0.0182	0.0000	0.0000	0.0000	0.0000	0.0091	0.0057	0.0041
All loci									
Na	8	11	14	12	6	12	7	7	12
H_e	0.6378	0.6168	0.6418	0.6521	0.5970	0.6261	0.6502	0.6361	0.6268
H_o	0.6703	0.6351	0.6280	0.6478	0.6036	0.6340	0.6746	0.6203	0.6425
F_{IS}	-0.0254	-0.0200	0.0266	0.0183	0.0397	-0.0023	-0.0029	0.0565	0.0325
Null	0.0250	0.0287	0.0206	0.0231	0.0308	0.0176	0.0202	0.0259	0.0240

Table 5. *S. zygaena* pairwise F-statistics and mitochondrial/nuclear ratios. Bold F-statistics are significant at $P < 0.000001$, italicized F-statistics are significant at $P < 0.05$, all others are non-significant. $mt\Phi_{ST}$ is based on mtCR haplotype frequencies and the genetic distance between haplotypes. mtF_{ST} is an analogue of $mt\Phi_{ST}$ that is based on mtCR haplotype frequencies. nF_{ST} is based on microsatellite frequencies. nF'_{ST} is standardized based on a recalculated maximum nF_{ST} value. Values of R are expected to vary between 1 and 4 depending on migration rates and time since population divergence. Collection areas are coded as in Table 1.

Population Pair	F-statistic				R (mitochondrial/nuclear ratio)			
Compared	$mt\Phi_{ST}$	mtF_{ST}	nF_{ST}	nF'_{ST}	$mt\Phi_{ST}$ / nF_{ST}	mtF_{ST} / nF_{ST}	$mt\Phi_{ST}$ / nF'_{ST}	mtF_{ST} / nF'_{ST}
Comparisons within Atlantic or IndoPacific								
WNA-WSA	**0.1116**	**0.6522**	0.0028	0.0077	39.73	232.18	14.54	85.00
WIO-WSP	**0.5290**	**0.4394**	*0.0022*	*0.0064*	239.05	198.54	83.05	68.97
WIO-WNP	**0.4864**	**0.3076**	0.0010	0.0028	475.04	300.36	173.48	109.69
WIO-ENP	**0.3581**	**0.4757**	0.0021	0.0060	167.27	222.18	59.90	79.56
WIO-ETP	**0.6524**	**0.3276**	0.0092	0.0266	70.97	35.64	24.52	12.31
WIO-ESP	**0.5838**	**0.3124**	-0.0006	-0.0016	1,048.59	561.11	370.89	198.46
WSP-WNP	**0.5351**	**0.5025**	0.0044	0.0125	121.29	113.90	42.89	40.27
WSP-ENP	**0.7702**	**0.6286**	0.0035	0.0101	217.57	177.56	76.43	62.38
WSP-ETP	**0.7462**	**0.5121**	0.0122	0.0366	61.21	42.01	20.37	13.98
WSP-ESP	**0.6778**	**0.4767**	-0.0055	-0.0160	123.65	86.95	42.25	29.71
WNP-ENP	**0.7625**	**0.5548**	0.0028	0.0074	277.15	201.68	103.35	75.20
WNP-ETP	**0.4632**	**0.3374**	-0.0045	-0.0129	102.36	74.57	35.90	26.15
WNP-ESP	**0.2912**	**0.3155**	0.0037	0.0102	79.55	86.20	28.62	31.01
ENP-ETP	**0.8345**	**0.5610**	0.0075	0.0213	111.46	74.94	39.12	26.30
ENP-ESP	**0.7856**	**0.5222**	-0.0001	-0.0001	14,642.68	9,733.27	5,264.04	3,499.11
ETP-ESP	**0.0949**	**0.3386**	0.0080	0.0237	11.88	42.35	4.01	14.28
Comparisons between Atlantic and IndoPacific								
WNA-WIO	**0.7857**	**0.4939**	**0.0476**	**0.1343**	16.50	10.37	5.85	3.68
WNA-WSP	**0.8852**	**0.6791**	**0.0553**	**0.1614**	16.00	12.28	5.48	4.21
WNA-WNP	**0.8761**	**0.6212**	**0.0361**	**0.0999**	24.29	17.22	8.77	6.22
WNA-ENP	**0.9067**	**0.7106**	**0.0519**	**0.1440**	17.48	13.70	6.30	4.93
WNA-ETP	**0.8735**	**0.6171**	**0.0373**	**0.1101**	23.40	16.53	7.93	5.60
WNA-ESP	**0.7596**	**0.5531**	**0.0577**	**0.1656**	13.17	9.59	4.59	3.34
WSA-WIO	**0.7697**	**0.4315**	**0.0458**	**0.1254**	16.81	9.43	6.14	3.44
WSA-WSP	**0.8469**	**0.5792**	**0.0625**	**0.1746**	13.54	9.26	4.85	3.32
WSA-WNP	**0.7667**	**0.4840**	**0.0449**	**0.1185**	17.08	10.78	6.47	4.09
WSA-ENP	**0.8727**	**0.6104**	**0.0545**	**0.1471**	16.00	11.19	5.93	4.15
WSA-ETP	**0.7553**	**0.4951**	**0.0468**	**0.1305**	16.12	10.57	5.79	3.79
WSA-ESP	**0.6991**	**0.4643**	**0.0636**	**0.1735**	10.99	7.30	4.03	2.68
Overall Atlantic/IndoPacific								
	0.8159	**0.6336**	**0.0495**	**0.3593**	16.48	12.80	2.27	1.76

Figures

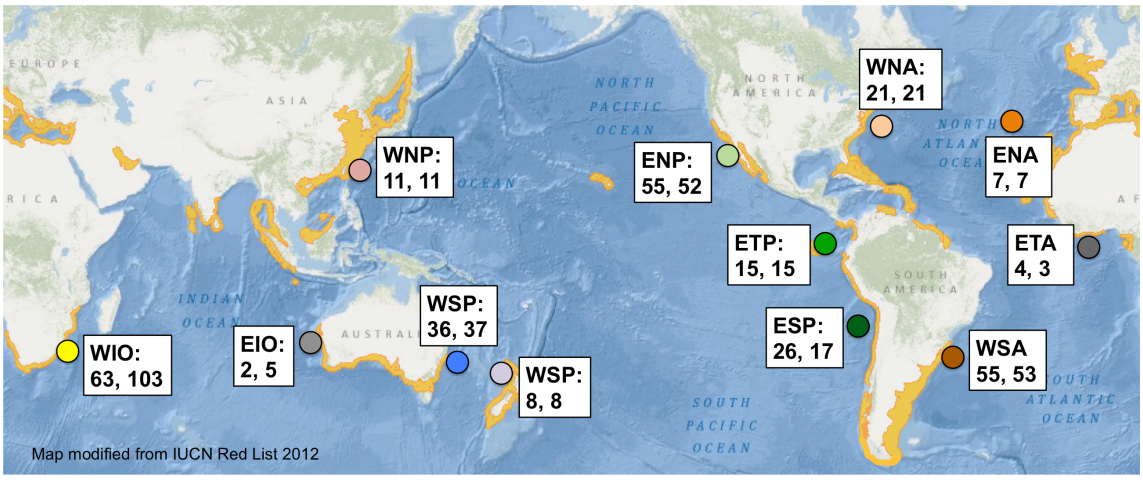


Figure 1. Map of *S. zygaena* distribution and sample sizes**.** Species distribution is shown in gold. Collection areas are represented by colored circles, the colors of collection areas correspond to colors used in Figures 2 and 3. Sample sizes are indicated in boxes, with the number of individual mtCR sequences first followed by the number of individuals genotyped. Collection areas are coded as follows: WIO, western Indian Ocean; EIO, eastern Indian Ocean; WSP, western South Pacific; WNP, western North Pacific; ENP, eastern North Pacific; ETP, eastern Tropical Pacific; ESP, eastern South Pacific; WNA, western North Atlantic; ENA, eastern North Atlantic; ETA, eastern tropical Atlantic; WSA, western South Atlantic.

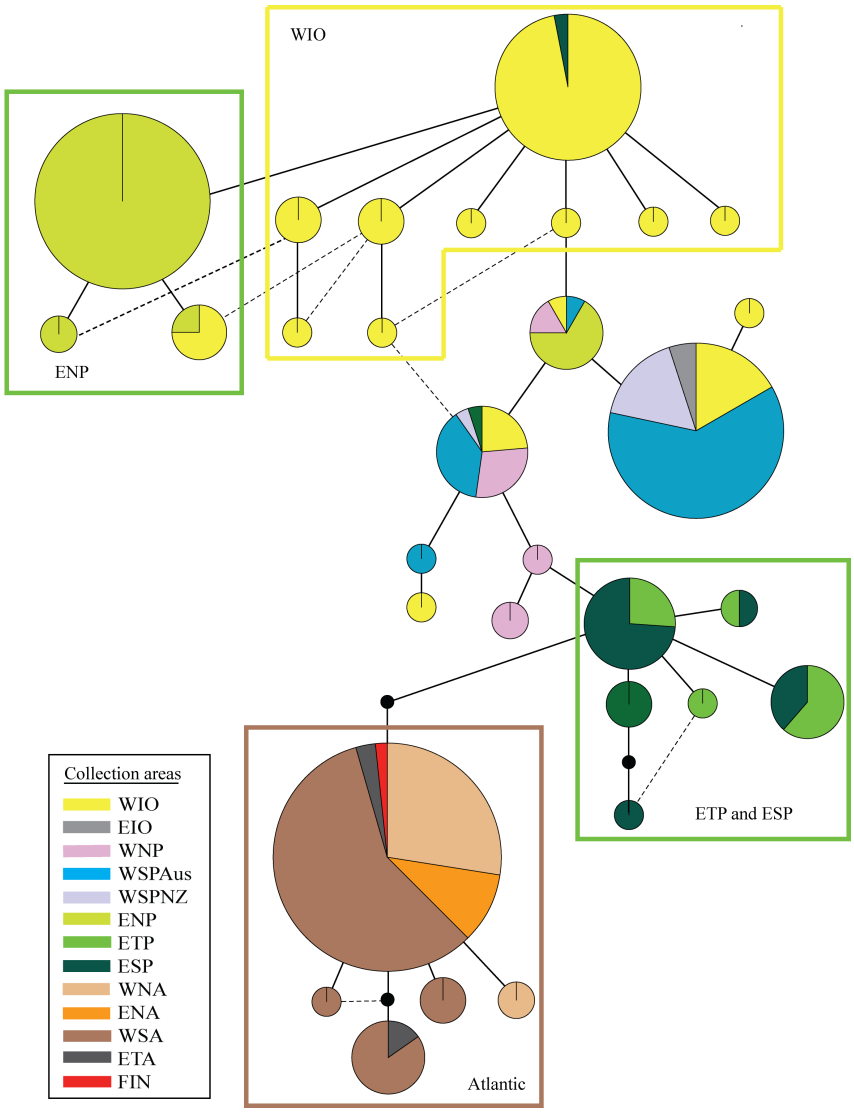


Figure 2. *Sphyrna zygaena* statistical parsimony haplotype network (mtCR). Circles represent individual haplotypes with circle size proportional to sampling frequency. Colored pie slices are proportional to the number of individuals from each collection area sharing that haplotype. Unbroken connecting lines are equivalent to one mutation step and small black circles represent inferred, unsampled haplotypes. Alternate connections are indicated as dashed lines. Shaded boxes indicate distinct geographic clusters. Collection areas are coded as in Figure 1.

Figure 3a. Principal coordinates (PCoA), mitochondrial control region sequences

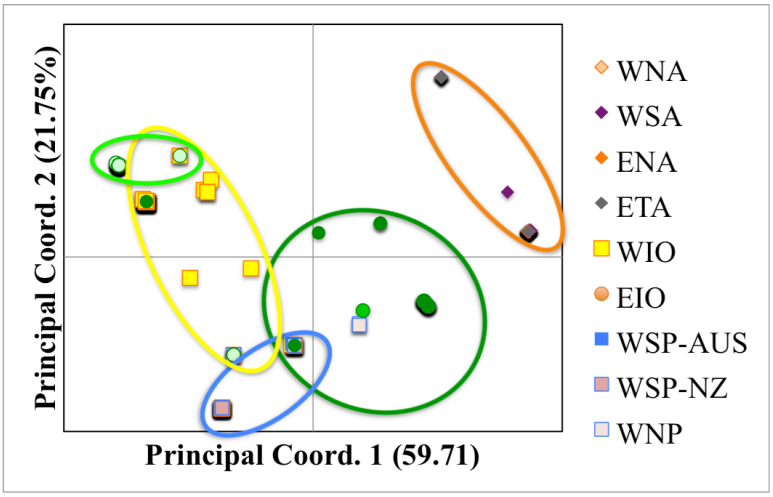


Figure 3b. Principal coordinates (PCoA), microsatellites

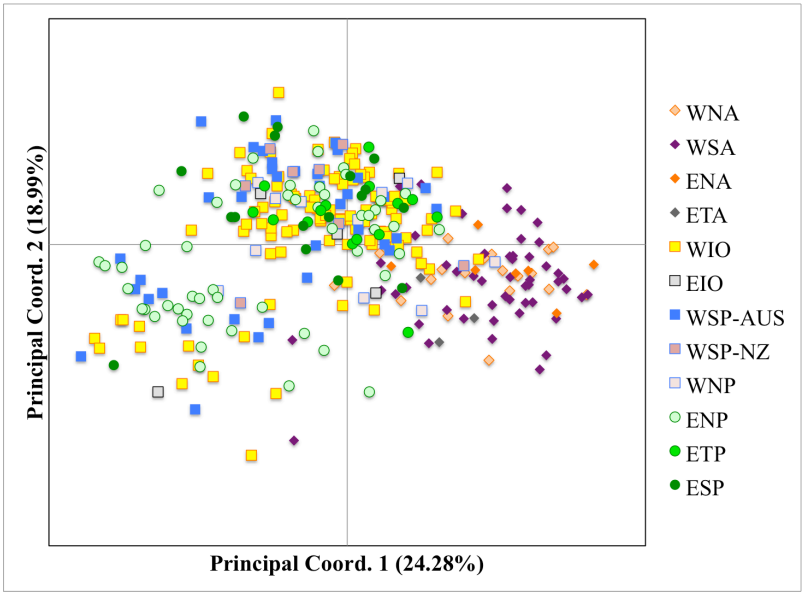


Figure 3. *S. zygaena* Principal Coordinate Analysis (PCoA) on (a) mitochondrial control region sequences and (b) nuclear microsatellite genotypes. Individuals from all collection areas were included, colored shapes correspond to collection areas according to the figure legends. The first two principal coordinate axes are shown with the amount of variance explained by each in parentheses. In Figure 3a, larger colored ovals indicate geographic clusters. Collection areas are coded as in Figure 1.

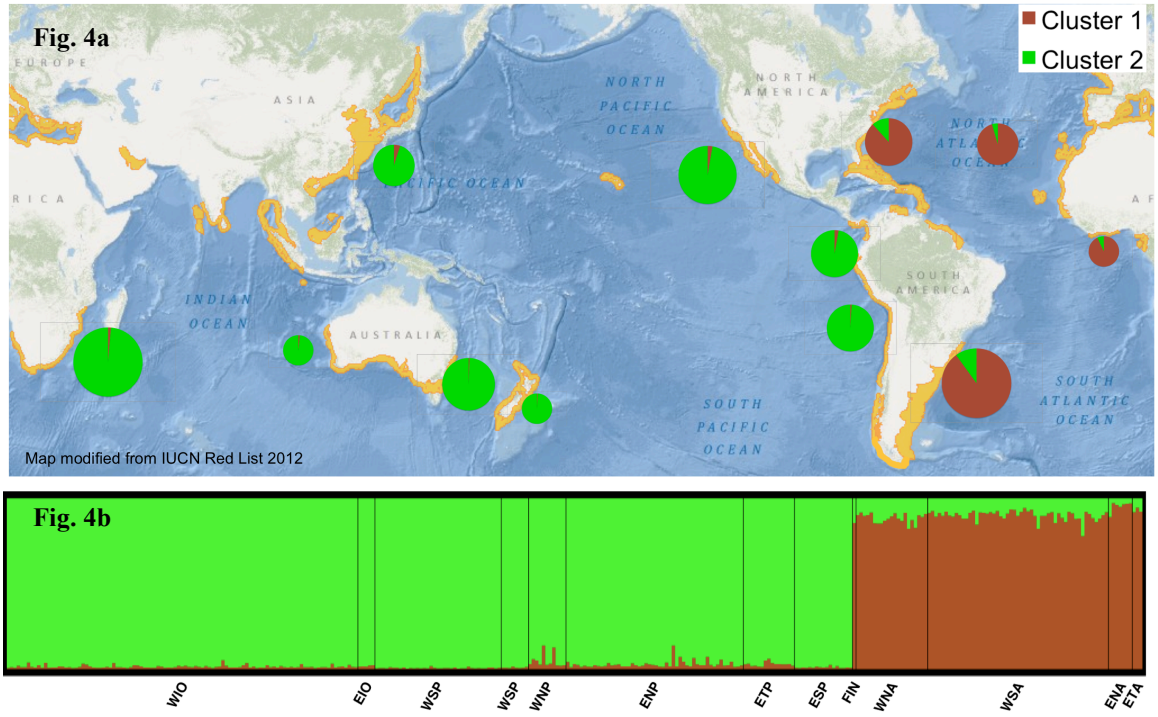


Fig 4. *S. zygaena* STRUCTURE analysis results (microsatellites). In Fig. 4a, pie charts indicate the average proportional membership coefficient of individual sharks in the two distinct clusters inferred from nuclear microsatellite genotypes by the program STRUCTURE. Pie chart sizes are proportional to sample sizes. Fig. 4b depicts assignment of individual sharks to each cluster in a conventional bar plot. Collection areas are coded as in Figure 1.

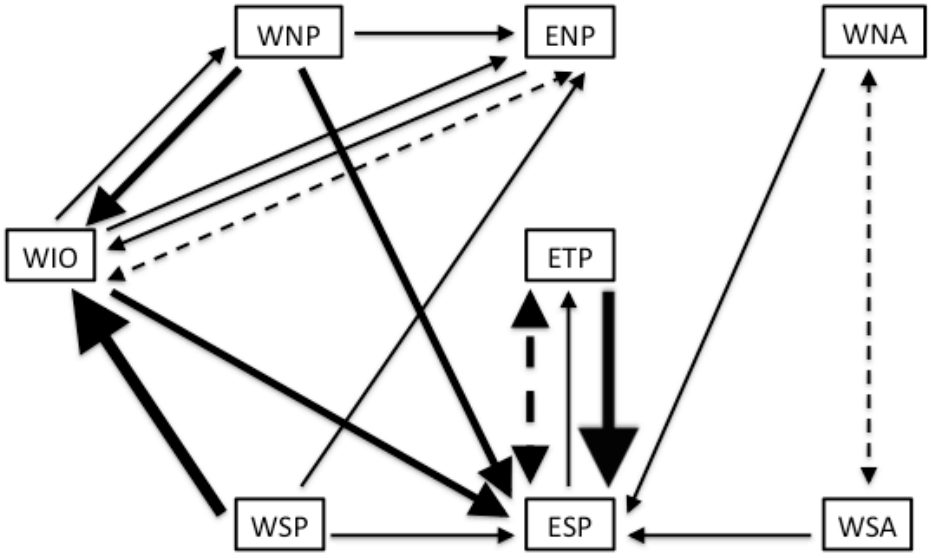


Fig. 5 Summary of the estimated gene flow among *S. zygaena* populations based on Bayesian inferences of migration rates and population sizes of mtCR using LAMARC v2.1.8 and MDIV. The arrows represent inferred migration direction, solid lines represent directional LAMARC results, double-headed arrows with dotted lines represent non-directional MDIV results, and thicknesses are proportional to $4N_em$. Collection areas are coded as in Figure 1.

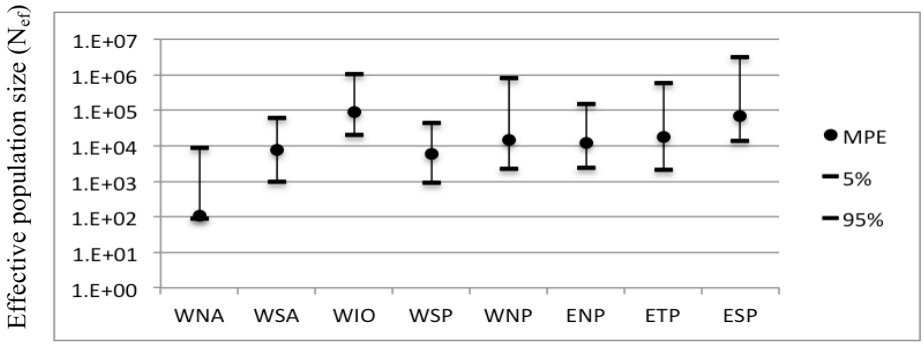


Figure 6a. Summary of the estimated effective population sizes (N_{ef}) and coalescent times of *S. zygaena* populations based on Bayesian inferences of mtCR data. The x-axis indicates the populations, with locations coded as in Figure 1. Figure 6a represents results from LAMARC, the y-axis is the effective population size (N_{ef}) in logarithmic scale, the central dot represents the most probable estimate (MPE), the vertical lines extend from the 5% to the 95% percentiles.

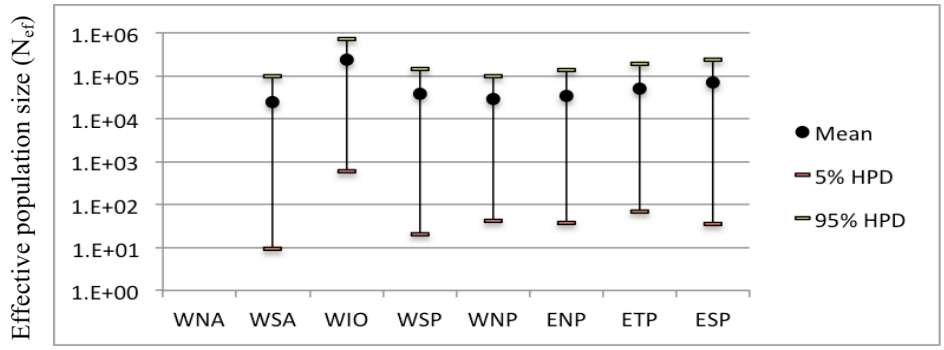


Figure 6b. Summary of the estimated effective population sizes (N_{ef}) and coalescent times of *S. zygaena* populations based on Bayesian inferences of mtCR data. The x-axis indicates the populations, with locations coded as in Figure 1. Figure 6b depicts results from BEAST, the y-axis is the effective population size (N_{ef}) in logarithmic scale, the central dot represents the mean, the vertical lines extend from the 5% to the 95% highest posterior distributions.

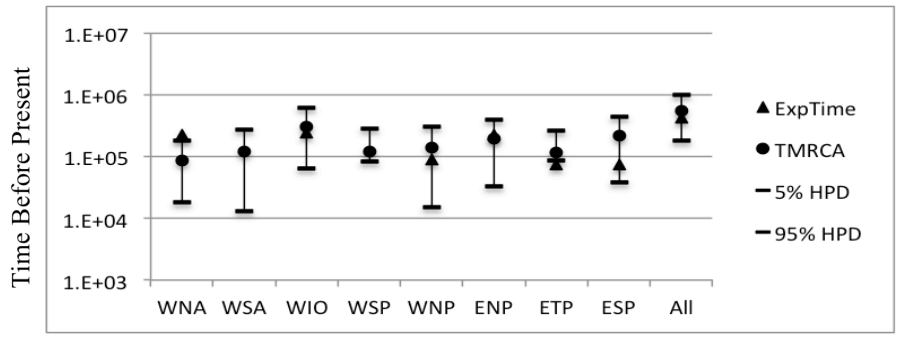
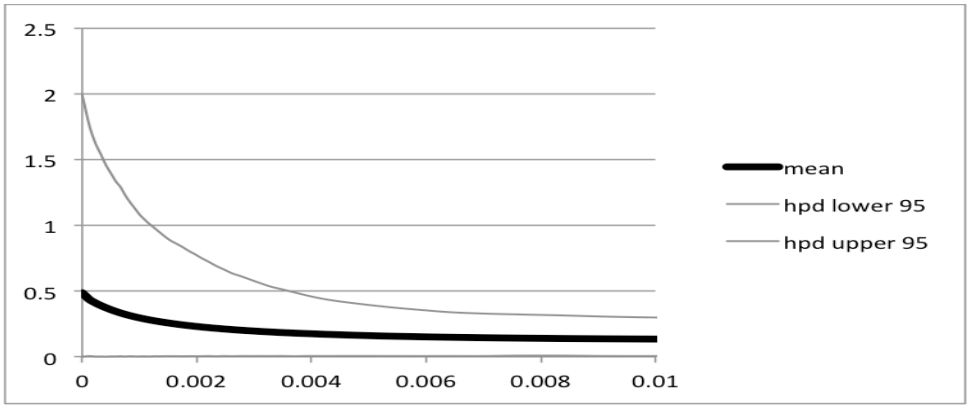


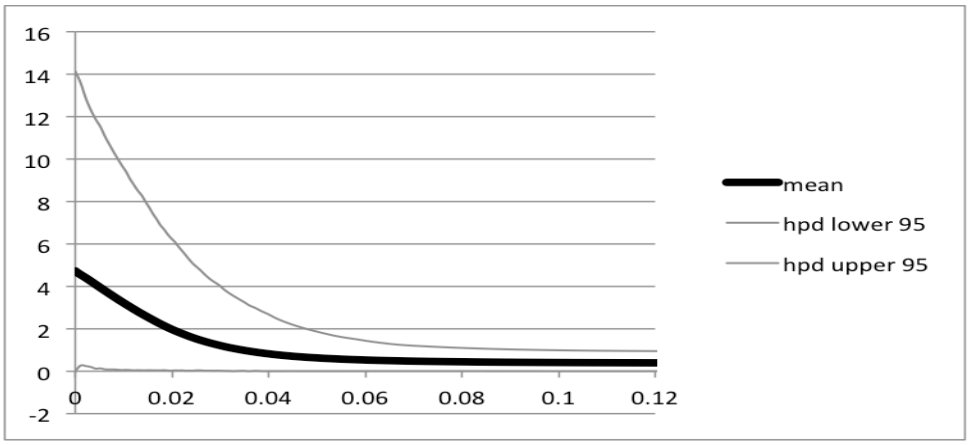
Figure 6c. Summary of the estimated effective population sizes (N_{ef}) and coalescent times of *S. zygaena* populations based on Bayesian inferences of mtCR data. The x-axis indicates the populations, with locations coded as in Figure 1. Figure 6c illustrates the time to most recent common ancestor (TMRCA) estimated in BEAST as the central dot, the expansion time calculated from τ in ARLEQUIN as a triangle, and the vertical lines extend from the 5% to the 95% TMRCA HPD. The y-axis is time before present in logarithmic scale.

Figure 7 – *S. zygaena* extended Bayesian skyline plots (mtCR and microsatellites)

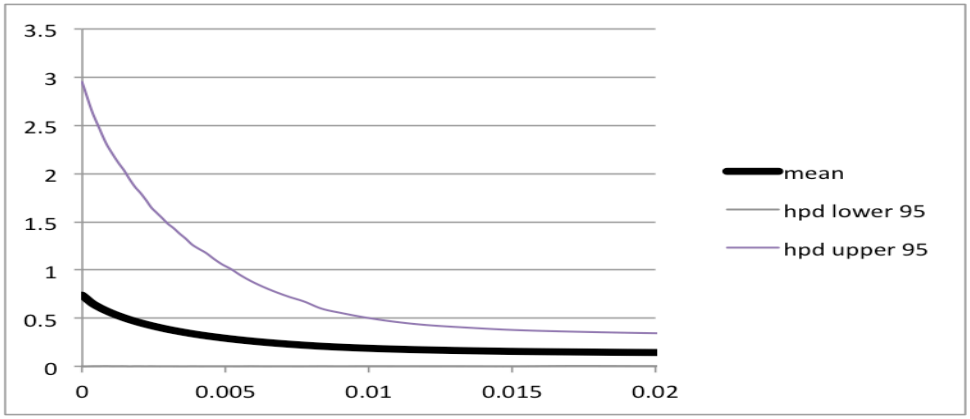
Western South Atlantic (mtCR)



Western Indian Ocean (mtCR)



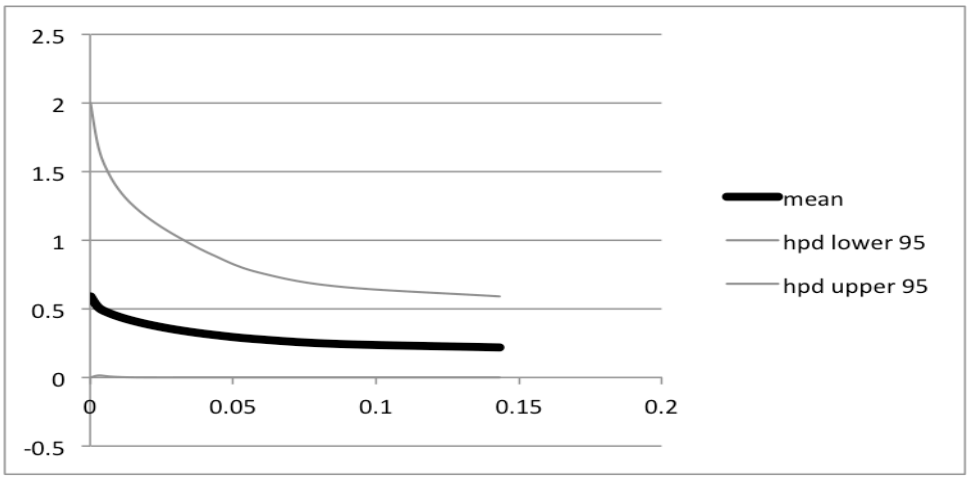
Western South Pacific (mtCR)



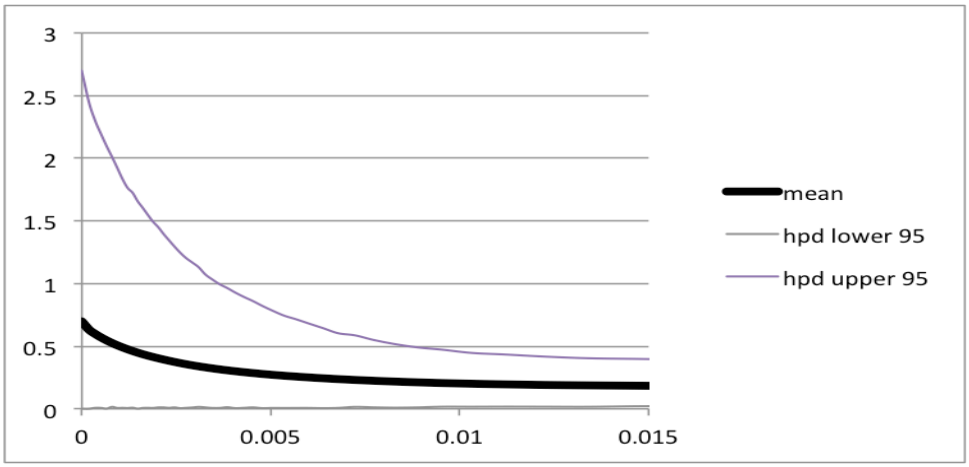
Effective population size (millions) times generation length

Time (millions of years before present)

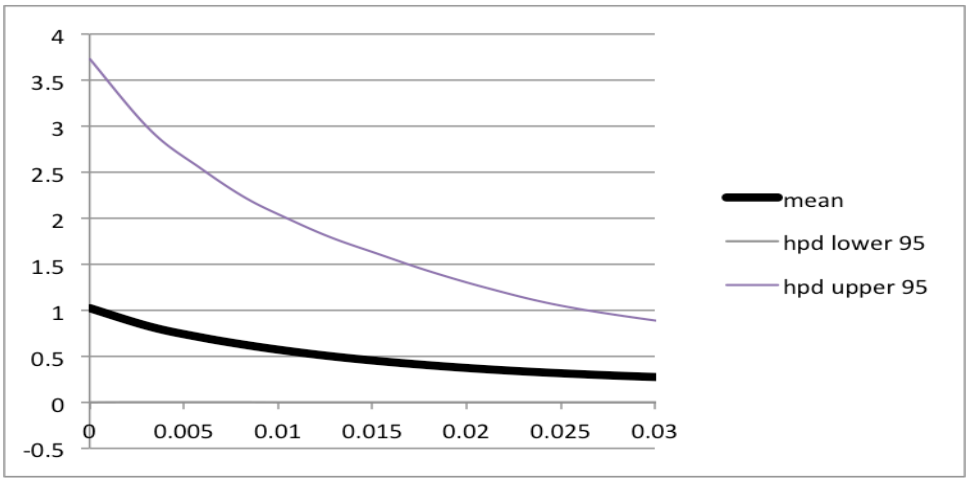
Western North Pacific (mtCR)



Eastern North Pacific (mtCR)



Eastern Tropical Pacific (mtCR)



Effective population size (millions) times generation length

Time (millions of years before present)

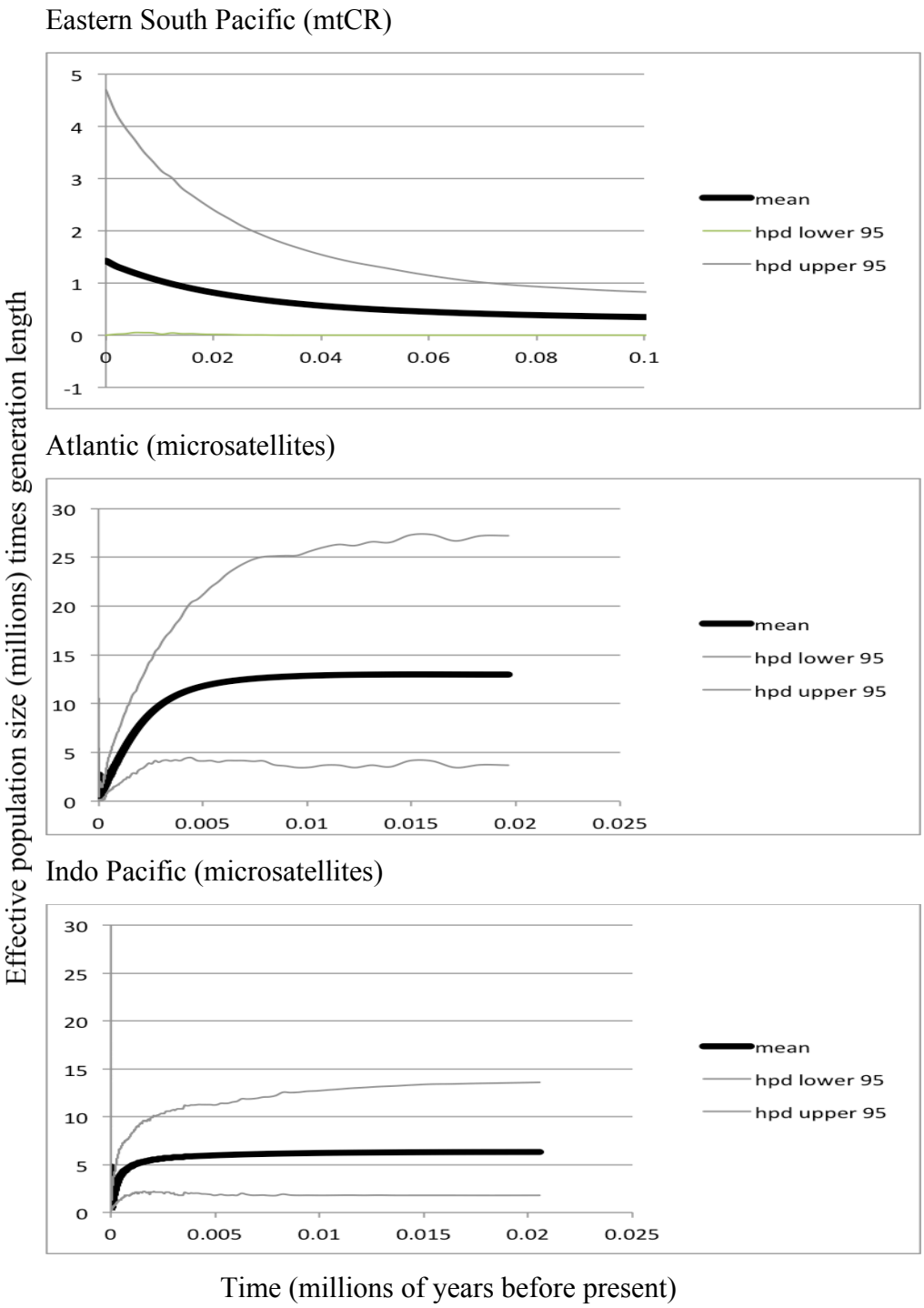


Fig 7. Extended Bayesian skyline plots for each genetically distinct *S. zygaena* population (mtCR and microsatellites). Solid black line shows mean effective size through time. Grey lines show the 95% highest posterior density.

Figure 8 – Summary of *S. zygaena* population sizes (microsatellites)

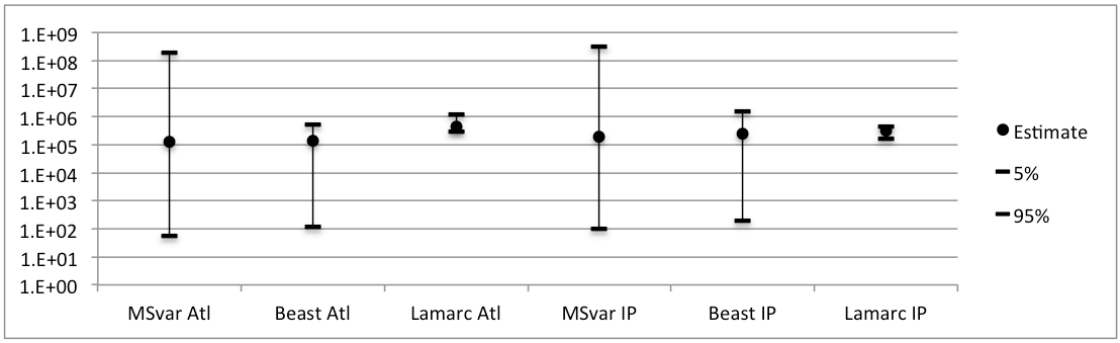


Fig. 8 Summary of the estimated effective population sizes (N_e) of *S. zygaena* populations based on Bayesian inferences of microsatellite data. The x-axis indicates the populations, either Atlantic or Pacific, and the analytical program used (MSVAR, BEAST, or LAMARC). The y-axis is the effective population size (N_e) in logarithmic scale, the central dot represents the point estimate (MSVAR - mode, BEAST - mean, LAMARC – most probable estimate (MPE)), the vertical lines extend from the 5% to the 95% highest posterior density (HPD) for MSVAR and BEAST, and percentiles for LAMARC.

Supplementary Information

Table S1. Microsatellite loci used to genotype *S. zygaena*

Locus	Annealing temperature (^{o}C)	Reference
Clim100	54	Keeney & Heist (2003)
Clim12	54	Keeney & Heist (2003)
Tgr47	57	Bernard *et al.* (in prep.)
Ghh-A6	55	Testerman *et al.* (in prep.)
Ghh-A8	59	Testerman *et al.* (in prep.)
Ghh-C7	55	Testerman *et al.* (in prep.)
Ghh-D1	59	Testerman *et al.* (in prep.)
Ghh-D10	50	Testerman *et al.* (in prep.)
Ghh-D9	52	Testerman *et al.* (in prep.)
Ghh-F1	52	Testerman *et al.* (in prep.)
Ghh-F3	52	Testerman *et al.* (in prep.)
Ghh-G5	50	Testerman *et al.* (in prep.)
SLE027	55	Nance *et al.* (2009)
SLE028	50	Nance *et al.* (2009)
SLE038	59	Nance *et al.* (2009)
SLE071	50	Nance *et al.* (2009)
SLE077	50	Nance *et al.* (2009)
SLE081	59	Nance *et al.* (2009)
SLE086	59	Nance *et al.* (2009)

Table S2. *S. zygaena* migration rates (mtCR). Migration rates are reported in terms of $4N_em$. Migration rate 1 ➔ 2 indicates the rate of migration from population 1 into 2, and vice versa. Collection areas are coded as in Table 1.

Population Pair Compared	Lamarc Results						MDIV Results			
	# migr per gen 1 -> 2			# migr per gen 2 -> 1						
	MPE	5% CI	95% CI	MPE	5% CI	95% CI	N_e	# migr per gen	Diverg. Time	TMRCA
WNA-WSA	0.78	0.00	2.92	0.01	0.00	0.03	27,014	2.21	14,588	1,133,509
WNA-WIO	0.45	0.00	2.34	0.01	0.00	0.01	15,873	0.01	59,046	94,284
WNA-WSP	0.15	0.00	0.91	0.00	0.00	0.02	3,663	0.00	49,449	52,306
WNA-WNP	0.37	0.00	2.21	0.01	0.00	0.02	3,587	0.00	27,473	39,955
WNA-ENP	0.18	0.00	1.09	0.01	0.00	0.01	4,884	0.00	61,244	71,989
WNA-ETP	0.42	0.00	2.55	0.01	0.00	0.02	3,663	0.00	27,765	40,732
WNA-ESP	1.35	0.00	7.50	0.01	0.00	0.01	8,928	0.01	39,642	64,820
WSA-WIO	0.66	0.00	4.48	0.18	0.00	0.69	13,965	0.01	63,959	98,871
WSA-WSP	0.16	0.00	0.87	0.20	0.00	1.12	4,502	0.00	5,223	6,303
WSA-WNP	0.30	0.00	1.61	0.23	0.00	1.29	4,960	0.00	34,424	51,388
WSA-ENP	0.20	0.00	1.30	0.21	0.00	1.06	5,342	0.00	61,537	76,814
WSA-ETP	0.44	0.00	2.45	0.20	0.00	1.05	4,808	0.00	32,595	49,999
WSA-ESP	1.01	0.00	6.86	0.23	0.00	0.72	7,860	0.00	39,614	62,723
WIO-WSP	0.00	0.00	0.39	25.82	9.45	26.76	13,278	0.70	29,477	60,814
WIO-WNP	1.31	0.00	5.27	8.79	0.00	10.14	19,459	0.83	26,465	71,999
WIO-ENP	1.15	0.13	1.24	1.34	0.00	6.12	15,262	1.09	19,230	68,375
WIO-ETP	0.44	0.00	1.12	0.70	0.00	4.28	19,993	0.05	47,185	78,774
WIO-ESP	3.59	0.42	3.86	0.00	0.00	0.66	21,596	0.40	44,488	81,201
WSP-WNP	0.90	0.00	2.62	0.72	0.00	0.78	4,350	0.11	5,829	30,796
WSP-ENP	1.25	0.00	2.10	0.06	0.00	0.22	4,731	0.02	29,145	52,422
WSP-ETP	0.44	0.00	2.50	0.19	0.00	0.89	5,571	0.00	25,291	45,234
WSP-ESP	2.37	0.00	5.45	0.28	0.00	1.47	9,463	0.11	16,276	58,100
WNP-ENP	1.13	0.00	1.71	0.18	0.00	0.58	5,876	0.04	38,664	57,702
WNP-ETP	0.85	0.00	4.75	0.98	0.00	2.48	6,944	0.11	14,305	38,888
WNP-ESP	6.33	0.00	7.21	0.60	0.00	2.92	11,141	0.41	18,272	52,142
ENP-ETP	0.17	0.00	1.02	0.62	0.00	1.81	6,029	0.00	30,143	59,080
ENP-ESP	0.83	0.00	4.55	0.43	0.00	1.54	9,463	0.05	22,142	66,616
ETP-ESP	13.66	0.00	14.84	2.92	0.00	14.44	9,539	5.11	267	46,931

Table S3. *S. zygaena* demographic analyses. Collection areas are coded as in Table 1.

Table S3a. Demographic analyses, mitochondrial control region. LAMARC: Reported values include effective population in millions (N_e), and exponential growth rate parameter (Population growth). Also reported are 95% confidence intervals. BEAST: Reported values inclue effective population size in millions (N_e) and time to most recent common ancestor in million years (TMRCA) with associated 95% highest posterior densities (HPD). ARLEQUIN: Time since population expansion in million years.

	Lamarc Results						Beast Results						Arlequin
	N_e (million)			Population growth			N_e (mil.)			TMRCA (MY)			Time Since
Population	MPE	5% CI	95% CI	MPE	5% CI	95% CI	mean	5% HPD	95% HPD	mean	5% HPD	95% HPD	Exp (τ)
WNA	107	88	8,927	6,481	-125	14,630	NA	NA	NA	85,925	18,176	184,000	224,809
WSA	7,745	1,000	62,572	1,227	-219	14,339	24,465	9	100,092	122,400	12,971	274,500	0
WIO	89,330	19,776	1,029,362	3,294	450	12,300	235,419	586	706,976	305,900	63,406	622,500	241,944
WSP	6,024	890	42,872	916	-306	13,937	37,148	20	147,712	121,700	83,583	282,700	0
WNP	14,987	2,277	787,799	2,697	-10	14,282	29,508	42	100,093	143,100	14,971	308,300	91,667
ENP	12,379	2,400	147,414	5,387	-173	13,797	34,940	37	134,986	194,100	32,546	395,700	227,462
ETP	17,883	2,097	586,769	6,869	262	14,581	51,075	70	186,123	115,400	85,898	265,000	75,063
ESP	71,332	13,555	3,127,927	7,635	541	14,339	71,457	35	234,918	218,100	38,480	452,400	74,759
Combined										549,000	184,900	1,016,600	424,748

Table S3b. Nuclear microsatellite tests of demographic change. MSVAR: Reported values include the current population size (N0), the historic population size (N1), the time in million years since demographic change, and the associated 95% HPD.

							MSvar Results		
	MSvar Results N_0			MSvar Results N_1			Time since event		
	mode	95% HPD		mode	95% HPD		mode	95% HPD	
Population									
Atl	1.E+05	6.E+01	2.E+08	9.E+06	4.E+03	1.E+10	4.E+06	2.E+03	9.E+09
IWP	2.E+05	1.E+02	3.E+08	8.E+06	5.E+03	1.E+10	5.E+06	3.E+03	9.E+09

Table S3c. Demograpic analyses, nuclear microsatellites. BEAST: Reported values inclue effective population size in millions (N_e) with associated 95% highest posterior densities (HPD). LAMARC: Reported values include effective population in millions (N_e), exponential growth rate parameter (Population growth), and the migration rates in terms of $4N_em$. Migration rate 1 ➔ 2 indicates the rate of migration from population 1 into 2, and vice versa. Also reported are 95% confidence intervals.

	Beast Results N_e			Lamarc Results N_e			Lamarc Results G			Lamarc Results			Lamarc Results		
	mean	5% HPD	95% HPD	MPE	5% CI	95% CI	MPE	5% CI	95% CI	# migr per gen 1 ➔ 2			# migr per gen 2 ➔ 1		
Population										MPE	5% CI	95% CI	MPE	5% CI	95% CI
Atl	135,638	120	524,953	196,301	123,791	511,420	0.00	0.53	1.48						
IWP	241,831	189	1,500,272	137,094	69,449	195,904	-2.27	-2.44	-0.54						
Comparison															
Atl-IWP										163	121	210	1.97	1.19	2.22

General Conclusions

Significant findings

This dissertation represents the first global survey of population genetic structure, genetic diversity, and evolutionary and demographic history of these four fisheries impacted sharks of conservation concern. This work was conducted to generate baseline genetic information that could be used to assist with conservation efforts.

Patterns of genetic differentiation and diversity varied among the species. There are two highly divergent genetic populations of porbeagle. Genetic diversity in the porbeagle is high compared to that reported for other sharks, and is largest in the North Atlantic. Interestingly, molecular data indicate a likely origin of the porbeagle in the North Atlantic with subsequent colonization of the Southern Hemisphere during Pleistocene glaciations. The bull shark exhibited no population differentiation along continuous coastlines, but strong differentiation across expanses of deep oceanic waters. There was strong population differentiation between great hammerhead samples from the western North Atlantic and Australia, although mito-nuclear discordance in the Mideast is possibly due to a combination of historic oceanic processes and more recent male mediated gene flow. Finally, the smooth hammerhead exhibits strong differentiation between the Atlantic and the Indo and Eastern Pacific oceans, as well as matrilineal genetic structure within oceanic basins, and an interesting dispersal pathway between the western Atlantic and eastern South Pacific.

Dispersal barriers can be quite different for different species. For example, warm equatorial waters are a barrier for the porbeagle. A lack of coastal connections makes the open ocean a barrier for the bull shark. Historic oceanographic processes appear to have shaped the current population structure of the great hammerhead, while female philopatry is the dominant factor in the smooth hammerhead. And finally, continental landmasses appear to influence connectivity in bull sharks and both hammerheads.

Although all four species can be partitioned into discrete genetic populations, in species with male mediated gene flow (i.e., the great hammerhead, smooth hammerhead, and likely the bull shark), the geographic boundaries of populations tend to be unclear because they differ for males and females. The protection of females at appropriate

spatial scales is critical to reduce the risk of localized extirpations due to philopatric behavior. Thus for conservation purposes these species should be managed at regional levels at the scale of the smallest identified population. Additionally, the genetic divergence between populations is large enough to develop forensic markers to track market products of all of these species, and a PCR based test of geographic origin was developed for the porbeagle shark.

Finally, the high levels of genetic diversity observed in these species suggest that the outlook for their survival is good, assuming that these measures reflect genome-wide diversity. However, much of the diversity was present as rare haplotypes or alleles, which could be highly susceptible to loss via genetic drift. This risk is likely to be mitigated at least to some extent by the relatively long lifespans of these species and the presence of overlapping generations.

Future directions

These general characterizations of population genetic structure set the stage for future studies to elucidate fine scale structure and explicitly test the hypotheses of sex biased dispersal and biogeographic boundaries proposed here. Specifically, future work on the porbeagle should include incorporation of nuclear markers and additional mitochondrial protein coding genes to rigorously test the hypothesis of speciation between the northern and southern populations. For bull sharks, additional samples are necessary to assess the genetic distinctiveness of other oceanic islands and the demographic history of the Eastern Pacific and Western Atlantic. For the great hammerhead, additional work would include incorporating samples to fully explore the patterns of mito-nuclear discordance and dispersal pathways. For the smooth hammerhead, more extensive sampling is necessary to elucidate fine scale regional patterns of connectivity and sex-biased dispersal. Population structure and genetic diversity can directly affect species survival. A thorough understanding of the molecular ecology of fisheries impacted species is crucial for their recovery.

General Acknowledgements

I am extremely grateful to my major advisor, Dr. Mahmood Shivji, for his tremendous support, inspiration, and advice. I also thank the rest of my committee for their guidance: Dr. George Duncan, Dr. Dave Gilliam, and Dr. Jose Lopez. My fellow lab mates, both present and past, have been a great source of encouragement, and advice, they are wonderful to work with and have made this long journey much more bearable: Debbie Abercrombie, Andrea Bernard, Derek Burkholder, Demian Chapman, Melissa Debiasse, Kim Finnegan, Teagen Gray, Marcy Henning-Plaza, Jennifer Hester, Rebekah Horn, Lucy Howey, Jennifer Magnussen, Lara Murphy, Alexandria Pickard, Vince Richards, Cassie Ruck, Shara Teter, and Jeremy Vaudo. This work could not have been done without the many people who graciously provided samples. Special thanks to the Guy Harvey Research Foundation, the Guy Harvey Ocean Foundation, the Save Our Seas Foundation, Sea Grant Florida, the Shark Foundation Hai-Stiftung, the Fish Florida Foundation, the Yamaha Contender Miami Billfish Tournament, the NSUOC SGA Student Travel Awards, and the American Elasmobranch Society Student Travel awards for their generous financial support. Dr. Richard Dodge, Dr. Richard Spieler, and the NSUOC faculty and staff, thank you for working so tirelessly to ensure the success of your students. My friends, both here in Fort Lauderdale and elsewhere, have been wonderfully encouraging and supportive. Finally, this dissertation is dedicated to my family: Mom; Dad and Lucia; Nick and Becki; my grandparents; my many aunts, uncles and cousins; and of course my husband Ted, his family, and our furry kids. You have all brought me love, laughter and joy through the years. Thank you!